VOL. XXIII, No. 5, May-June, 1950.

MATHEMATICS

magazine

MATHEMATICS MAGAZINE

Formerly National Mathematics Magazine, founded by S. T. Sanders.

Address editorial correspondence to Glenn James, special papers to the editors of the departments for which they are intended, and general papers to any member of the Executive Committee.

Manuscripts should be typed on 8½" x 11" paper, double-spaced with 1" margins. We prefer that, in technical papers, the usual introduction be preceded by a forward which states in simple terms what the paper is about.

The Mathematics Magazine is published at Pacoima, California by the managing editor, bi-monthly except July-August. Ordinary subscriptions are $3.00, sponsoring subscriptions $10.00, single copies 65¢. Reprints: 25 free to authors; 100 @ $1.25 per page with a minimum of $6.00. All reprints are bound.

Subscriptions and related correspondence should be sent to Inez James, 14068 Van Nuys Blvd., Pacoima, Calif.

Advertisers should contact Homer V. Craig, Univ., of Texas, Austin 12, Tex.

Entered as second-class matter March 23, 1948 at the Post Office, Pacoima, California under act of congress of March 3, 1876.

CONTENTS

NOTE ON CONVERGENCE IN TOPOLOGY

Richard Arens

FOREWORD

Topology, which is a generalization of geometry, axiomatizes the notions of open set and closed set. Here we show a method of arriving at the same result by axiomatizing the notion of convergence. To understand what the difference in these approaches is, one should have had a course in point set topology. However to see wherein this paper differs from 10 or 20 others on the same topic (which we explain in the introduction) one should be acquainted with some of the references in the bibliography.

---

1. One of the earlier activities in the field of point-set theoretic topology was the study of "convergence spaces" in which there was defined a mode or notion of convergence of sequences of points. The axiomatic definition of these notions varied in content and appearance to suit the purpose and style of various authors (*cf.* the relevant sections of the texts listed at the end of this note). Partly for this reason, point set theory is now more commonly based on the notions of "open set" or of "closure" than on convergence. More recently, the class of possible subjects for the verb 'to converge' has been greatly increased (Birkhoff III; Tukey; Cartan (see Bourbaki)), and it has been shown that the topology of a space can be characterized by stating which of the possibly convergent things do converge and if so, to what. The results of the present note deal with a case in which all the different "generalized sequences $\{x_n\}$" involved are based on the *same system* of subscribed indices. In particular, the system of indices may be the class of natural numbers. Thus we have here a portion of the theory of those spaces in which limits are always accessible through ordinary sequences.

The paper had its origin in reflections on the notion of "convergence in measure" and as to why it is brought into the study of integration when at first glance "convergence almost everywhere" seems to be the desired phenomenon. The answer (considered in detail at the end of the paper) is that both these convergences determine the same topology, but this topology determines only convergence in measure. The ideas here presented are related to those expressed by G. Birkhoff (IV, p. 30) concerning "star-convergence".

2. A *partially ordered* set D is one in which a binary relation $m<n$ or $n>m$ is defined such that $m>p$ if $m>n$ and $n>p$. D is moreover *directed* if for $m, n \in D$ there is some $p>m, n$. We shall always use 'ω' to designate the directed set of natural numbers $\{1, 2, \cdots\}$, which is the chief example.

Let D be partially ordered and let f be a function defined on D with values in D, i.e., $f(n) \varepsilon D$ for $n \varepsilon D$. We shall say that f *has cofinal*

values if for each $p \in D$ there is an $n \in D$ such that $f(n) > p$. We shall say that $f(n) \to \infty$ if for each $p \in D$ there is an $m \in D$ such that $f(n) > p$ for all $n > m$.

Let X by any class, and let D be any directed set, and let ξ be a function defined on D with values in X, i.e., $\xi(n) \in X$ for $n \in D$. Then we call ξ a *system of points in X directed by D.*[1] The set of values $\{\xi(n)\}$ is called the *range* of ξ. The directed system ξ is of course not the same thing as its range. A system of points directed by the directed set ω of natural numbers is called a *sequence*. One example should reconcile the reader to this terminology if it is new to him: consider the sequence

2.1 $$(5,\ \frac{1}{2},\ 0,\ 3,\ -1,\ 4,\ \cdots).$$

It is not stretching the meaning of 2.1 to say that it means merely that $\xi(1) = 5$, $\xi(2) = \frac{1}{2}$, $\xi(3) = 0$, $\xi(4) = -1$, $\xi(5) = 4$, *etc*. This sequence is distinct from

2.2 $$\eta = (\frac{1}{2},\ 5,\ 0,\ 3,\ -1,\ 4,\ *,\ * \cdots) \quad (* = \text{same as for } \xi).$$

for $\eta(1) = \frac{1}{2} \neq 5 = \xi(1)$. However, they have the same range, *viz.*,

2.3 $$\{\frac{1}{2},\ 5,\ 0,\ 3,\ -1,\ \cdots\} = \{5,\ \frac{1}{2},\ 0,\ 3,\ -1,\ \cdots\}.$$

Note that if ξ is a system directed by D and if f is a function defined on D with values in D and we define

2.4 $$\eta(n) = \xi(f(n)) \in X,$$

then η is also a system directed by D. If $f(n) \to \infty$, then η is called a *directed subsystem* of ξ(*Cf.* " subsequence").

3. Let D and X be as above.

A *limit process 'lm' based on D and X* is a law which assigns to each system ξ of points in X directed by D[2] a class of points of X. If x is in the set of points which *lm* assigns to ξ, we shall write

3.1 $$\mathrm{lm}_n\ \xi(n) = x.$$

With a given , there may be several, only one, or no points x for which 3.1 holds.

We present now some axioms which some directed systems satisfy. Henceforth m, n, p will denote elements of the directed set D; $f, g, h, \cdots$

[1] In this context it is customary to write ξ_n instead of $\xi(n)$, but there is no logical reason for departing from the standard functional notation.

[2] We shall abbreviate this to "directed system ξ".

will denote functions defined on D with values in D; and ξ, η, possibly with indices, will denote directed systems.

3.2 If $\xi(n) = x$ for every n, then $\mathrm{lm}_n \xi(n) = x$.

3.3 If $\mathrm{lm}_n \xi(n) = x$ and if $f(n) > n$ for each n then $\mathrm{lm}_n \xi(f(n)) = x$.

3.4 Suppose $\mathrm{lm}_n \xi_m(n) = x_m$ for each $m \in D$, and $\mathrm{lm}_m x_m = x$.[1] Then there are f and g such that $\mathrm{lm}_n \xi_{f(n)}(g(n)) = x$.

3.41 Suppose $\mathrm{lm}_n \xi_m(n) = x_m$ for each m, and $\mathrm{lm}_m x_m = x$. Then there are f and g such that $f(n) \to \infty$ and $\mathrm{lm}_n \xi_{f(n)}(g(n)) = x$.

3.5 Let x and ξ be such that whenever $f(n)$ has cofinal values a g can be found such that $\mathrm{lm}_n \xi(f(g(n))) = x$. Then $\mathrm{lm}_n \xi(n) = x$.

We remark two things. If '$f(n) > n$' in 3.3 is replaced by '$f(n) \to \infty$,' a stronger axiom is obtained. Hence we have chosen 3.3. Also, 3.41 is stronger than 3.4.

4. For reference, we enumerate the axioms of a *topological space* X, taking as primitive notion that of closure (see Alexandroff-Hopf, *Topologie* I, p. 37, Kuratowski's axioms).

4.1 $$(A \cup B)^- = A^- \cup B^- \qquad (A, B \text{ subsets of } X)$$

4.2 $$A \subset A^-$$

4.3 $$(A^-)^- = A^-$$

4.4 $$\Lambda^- = \Lambda \qquad (\Lambda \text{ is the void set}).$$

If $A = A^-$, A is *closed*; if the complement of A is closed, A is *open*. In terms of open sets, $x \in A^-$ if and only if every open V which contains x intersects A in a non-void set.

To be definite, we quote the following definition:

4.5 A system ξ of points of a topological space X directed by D *converges* to a point $x \in X$ if and only if for every open set V containing x there is an $m \in D$ such that $\xi(n) \in V$ for all $n > m$. If ξ converges to x, we shall write

4.6 $$\lim_n \xi(n) = x.$$

The limit process defined by 4.6 satisfies 3.2 and 3.3, but it may not satisfy 3.4 or even 3.41 as we shall see later.

We wish to give a characterization of 4.6 in terms of closure.

4.7 *Theorem: 4.6 holds if and only if $x \in \{\xi(f(n))\}^-$ whenever f has cofinal values.*

Proof: First suppose $\lim_n \xi(n) = x$, but that $x \notin \{\xi(f(n))\}^-$. By 4.5,

[1] This latter really means that if we define $\eta(m) = x_m$ then $\mathrm{lm}_m \eta(m) = x$. Hence x_m may be regarded as an abbreviation for $\eta(m)$ where η is the directed system which x_m defines. The expression $\xi_{f(n)}(g(n))$ has a similar interpretation.

there exists an $m \in D$ such that $\xi(p) \notin \{\xi(f(n))\}$ for $p > m$. Hence $f(p)$ is never $> m$. Hence f does not assume cofinal values. Conversely let $\xi(n)$ not converge to x. Then by 4.5 there is an open V containing x such that for each m there is an $n = f(m) > m$ such that $\xi(n) \in V$. Clearly f assumes cofinal values but since V does not intersect $\{\xi(f(n))\}$, $x \notin \{\xi(f(n))\}^-$.

5. The first result here relating an abstractly given limit process to topological spaces is as follows.

5.1 *Theorem: Let the limit process* lm *based on X and D satisfy 3.2-3.4. For each subset A of X define A^- as follows:*

5.2 A^- is the class of all x for which ξ can be found with $\xi(n) \in A$ such that $\mathrm{lm}_n \xi(n) = x$.

Then the operation '$^-$' *satisfies* 4.1-4.4 *and* $\mathrm{lm}_n \xi(n) = x$ *implies* $\lim_n \xi(n) = x$ *(see* 4.6*).*

Proof: 4.2 and 4.4 are obvious; and since 5.2 entails $A^- \subset C^-$ when $A \subset C$, we need in 4.1 only show $(A \cup B)^- \subset A^- \cup B^-$. Now if $x \in (A \cup B)^-$ then either there exists an m such that $\xi(n) \in B$ for all $n > m$ or for every m there is an $n = f(m) > m$ with $\xi(f(m)) \in A$. In the former case $x \in B^-$ and in the latter case $x \in A^-$, by 3.3. For 4.3 let $x \in (A^-)^-$. Then there exists a ξ such that $\xi(n) \in A^-$ and $\mathrm{lm}_n \xi(n) = x$. Moreover, for each n there exists a ξ_n such that $\xi_n(m) \in A$ and $\mathrm{lm}_m \xi_n(m) = \xi(n)$ for each n. An application of 3.4 yields $x \in A^-$. Now suppose $\mathrm{lm}_n \xi(n) = x$. Let f have cofinal values. For each m there is an n with $f(n) > m$. Let $g(m) = f(n)$. Then $\mathrm{lm}_m \xi(g(m)) = x$ by 3.3, whence $x \in \{\xi(f(n))\}^-$. By 4.7, $\lim_n \xi(n) = x$. This completes the proof of 5.1.

5.3 *Corollary: Let* 3.4 *in* 5.1 *be replaced by* 3.41. *Then a directed set ξ which converges to x has a subsystem $\xi(f(n))$ for which* $\mathrm{lm}_n \xi(f(n)) = x$.

Proof: Let $\lim_n \xi(n) = x$. For each m, $n \in D$ define $f_m(n) > m$, n. By 4.7, $x \in \{\xi(f_m(n))\}^-$. By 5.2, there exists a g_m with $g_m(n) = f_m(p) > m$ for some suitable value of p with $\mathrm{lm}_n \xi(g_m(n)) = x$. On the other hand $\mathrm{lm}_m x = x$. Hence by 3.41, there exist h and k, $h(n) \to \infty$, such that

$$\mathrm{lm}_n \xi(g_{h(n)}(k(n))) = x.$$

Therefore, let

$$l(n) = g_{h(n)}(k(n)).$$

Then $l(n) > h(n) \to \infty$, and thus $\eta(n) = \xi(l(n))$ is a directed subsystem of ξ which has $\mathrm{lm}_n \eta(n) = x$, as desired.

The hypotheses of 5.1 and 5.3 do not suffice to prove that $\lim_n \xi(n) = x$ implies $\mathrm{lm}_n \xi(n) = x$, as the following example shows.

Let X = the real number system, and let $D=\omega$. For each sequence x_n of real numbers write $\mathrm{lm}\, x_n = x$ when $\Sigma|x_n - x| <\infty$. Now $\lim 1/n = 0$, but $\mathrm{lm}_n(1/n)=0$ is false. We leave the study of this example to the reader, and turn to a condition which enables us to conclude $\mathrm{lm}_n\xi(n)=x$ when $\lim_n \xi(n) = x$.

5.4 *Corollary. Adjoin* 3.5 *to the hypothesis of* 5.1. *Then* $\lim_n\xi(n) = x$ *if (and only if)* $\mathrm{lm}_n\xi(n) = x$.

Proof: Let $\lim_n \xi(n) = x$. Let $f(n)$ have cofinal values. Then $x\in\{\xi(f(n))\}^-$, and so g can be found (by 5.2) so that $\mathrm{lm}_n\xi(f(g(n)))=x$. By 3.5 $\mathrm{lm}_n\xi(n) = x$, as desired.

6. This section is devoted to some examples. D shall always be ω, the directed set of natural numbers.

The first example X is a topological space, in which 6.1 every point of A^- is the limit of a sequence in A $(A\subset X)$, but in which the first axiom of countability [Hausdorff, p. 229 (9)] is not fulfilled.

X consists of all pairs (m,n) of natural numbers and one "ideal" point z. We refer to the class of points (m,n) (n variable, m fixed) as the m-th *column*, $m=1,2,\cdots$. For $A\subset X$ we define

$$A^- = \begin{cases} A \text{ if } A \text{ has only finitely many points on each column} \\ A\cup\{z\} \text{ otherwise}\end{cases}$$

The reader may verify 4.1-4.4. In fact, X is a zero dimensional completely normal [Alexandroff-Hopf, p. 69] space.

Let us reformulate the first axiom of countability as follows.

6.2 *For each* $x\in X$ *there exists a countable family of closed sets* $\{A_1, A_2, \cdots, A_n, \cdots\}$ *none containing* x *and such that if any closed set* A *does not contain* x *then* A_n *contains* A *for at least one* n.

Now 6.2 is *violated* for $x=z$ in our example. Each A_n is closed and hence meets the n-th column in a *finite* set F_n. Select a finite subset G_n on the n-th column such that $G_n\supset F_n$ but $G_n\neq F_n$, and let $A = G_1\cup G_2\cup\cdots\cup G_n\cup\cdots$. Now A is not contained in any A_n.

Now when can $\xi(n)$ converge? If and only if $\xi(n)$ ultimately decides to lie on *one* column, and on that column becomes ultimately constant or tends upward. Thus 5.1 and 5.4 apply. Hence we have shown 6.1.

Now make the *set* X and adjoin points "1", "2", $\cdots$, "n", $\cdots$; where n shall be called '*the capital*' of the n-th column. For $A\subset Y$ define $A^- = A\cup B\cup C$ where B is the class of capitals of those columns which A intersects in infinite sets, and where C is void if B is finite and has z as its sole member if B is infinite. This Y is easily seen to be also a zero-dimensional completely normal topological space. The topological limit process *lim* here satisfies 3.2, 3.3, 3.5, but not 3.4. To see that 3.4 fails, take $\xi_m(n) = (n,m)$. Then $\lim_n\xi_m(n) = m$ (a capital) and $\lim_m m = z$. But 3.4 fails because z is the limit of no sequence from the

union of the columns. To see this, observe that when a sequence converges, in a Hausdorff space, its limit is also its only limit-point. Hence the sequence presumably converging to z could have only finitely many elements on each column (otherwise that capital would be a limit-point) whence not even z would be a limit-point.

A fortiori, this Y cannot satisfy the first axiom of countability.

Let Z be the complement of the class of capitals in Y, and for $A \subset Z$ define $A^{-} = Z \cap$ (the closure of A in Y). This makes Z into a a non-discrete topological space in which all compact sets are *finite*. Also, all compact sets are *open*. No sequence from $Z - \{z\}$ can converge to z as before.

Finally, consider the class S of finite-valued, measurable functions on a set of finite measure (*cf.* E. J. McShane, pp. 160 *et seqq.*). The theorems on term-by-term integration peculiar to Lebesque integration leave no doubt as to the importance of the limit process $\mathrm{lm}_n f_n = f$ which means that $f_n(x)$ approaches $f(x)$ for almost all x ("convergence almost everywhere"). The relations between convergence almost everywhere and convergence in measure (*cf.* McShane) make it clear that 5.1 and 5.3 apply and that convergence in measure is merely the *lim* of 4.6 in the topological space S in which closure has been defined (in 5.2) by means of *lm*, i.e., convergence almost everywhere. Stated in another way, convergence almost everywhere and convergence in measure determine the same topological space S, but (in general) only convergence in measure satisfies 3.5, and hence also the conditions of Corollary 5.4.

REFERENCES

C. Kuratowski, Topologie I, Warzawa (1933).

G. Birkhoff, I Axiomatics of Perfectly separable metric spaces, Bull. Am. Math. Soc. 39, 601-607 (1933).

II Topology of Transformation sets, Ann. of Math. 35, 861-875 (1934).

III Moore-Smith convergence in general topology, Ann. of Math. 38, 39-56 (1937).

IV *Lattice Theory* New York (1940 or 1948).

J. W. Tukey, Convergence and Uniformity in Topology, Princeton (1940).

N. Bourbaki, Elements de Mathematique II, premiere partie, Les Structures fondamentales de l'analyse, Livre III, Topologie Generale.

F. Hausdorff, Mengenlehre, (Dritte Auflage), Berlin (1935).

P. Alexandroff und H. Hopf, Topologie I. Berlin (1935)

E. J. McShane, Integration, Princeton (1947).

University of California at Los Angeles.

COLLEGIATE ARTICLES

Graduate Training not required for Reading

DUALITY AND DIFFERENTIAL CALCULUS IN THE PLANE

R. T. Hood

Early in the nineteenth century the fast-developing field of projective geometry was enriched by the introduction of a notion known as the Principle of Duality, which states that any theorem in projective geometry dealing with points and lines can give rise to another theorem simply by substituting lines for points in it, and vice versa. The one theorem is called the dual of the other. This principle was apparently recognized by Poncelet and used by him extensively, although it was not stated as an independent principle until 1826 (by Gergonne).

This principle implied that in projective geometry the point and the line were abstractly identical. A consequence of this was that geometers began to realize that it was no longer necessary to consider the point as the fundamental element of the plane. One could equally well start with a line. Then instead of a line being thought of as the totality of points lying on it, a point could be thought of as the totality of lines passing through it. A curve could be thought of as made up of its tangents instead of its points. It was thus possible to build up a whole new geometry based upon the line as the fundamental element. Because of the importance of tangents in this geometry, it was called tangential or line geometry as opposed to the cartesian or point geometry with which we are so much more familiar.

In analytic tangential geometry it would seem reasonable to have a coordinate system in which lines are represented by coordinates and points and curves by equations. Möbius introduced such a system in 1827. The most satisfactory set of coordinates to use is the set of negative reciprocals of the intercepts of the line in question. These will be denoted by u and v, respectively. Plücker used these coordinates in his work in 1829. All lines except lines passing through the origin can be represented by them, just as in point geometry all points can be represented by cartesian coordinates except those situated on the line at infinity. There is thus seen a dual correspondence between the origin and the line at infinity.

Any line not passing through the origin has an equation of the form

$$\frac{x}{a} + \frac{y}{b} = 1,$$

where a and b are its intercepts, or

$$(1) \qquad \left(-\frac{1}{a}\right)x + \left(-\frac{1}{b}\right)y + 1 = 0.$$

The coordinates of the line are $u = -\frac{1}{a}$, $v = -\frac{1}{b}$; that is, they are the coefficients of x and y, respectively in (1).

Consider a line with coordinates (u_1, v_1). Its point equation is

$$(2) \qquad u_1x + v_1y + 1 = 0.$$

If the point (x_1, y_1) lies on this line, then

$$u_1x_1 + v_1y_1 + 1 = 0.$$

Now let u_1 and v_1 become variables. Writing them u and v, we have the equation

$$(3) \qquad x_1u + y_1v + 1 = 0.$$

Every pair of values u and v satisfying this equation represents a line which passes through the point (x_1, y_1). Equation (3) represents a family of lines through (x_1, y_1). In line geometry a point is considered as a pencil of lines. Therefore, (3) is the line equation of the point (x_1, y_1).

The line equation of the point (x_1, y_1) is $x_1u + y_1v + 1 = 0$.

The point equation of the line (u_1, v_1) is $u_1x + v_1y + 1 = 0$.

A whole analytic geometry, based on line coordinates u and v instead of point coordinates x and y, has been developed. One result from this geometry is necessary to the later portion of this paper — the equation of the point of intersection of two lines (u_1, v_1) and (u_2, v_2). It is

$$(4) \qquad \begin{vmatrix} u & v & 1 \\ u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \end{vmatrix} = 0.$$

Consider the equation

$$(5) \qquad v = f(u),$$

expressing v as a single-valued continuous function of u for at least some range of values of u. For every value of u in that range, equation (5) determines a corresponding value of v, thus designating a certain line in the plane. As u varies continuously, v varies continuously, and the line moves continuously, enveloping a curve as it moves. Equation (5) is said to be the equation of this curve.

It is the purpose of the remainder of this paper to investigate the derivative $f'(u)$ in line geometry (this will be called the *line derivative* as opposed to $\frac{dy}{dx}$ which will be called the point derivative),

to see what uses can be made of it, and to call attention to certain dual facts and procedures which may be uncovered in the process. Let a curve (in line geometry this may mean a point, since a point also has an equation) have the equation (5) restricted as above. Then the line derivative is defined as one would expect,

$$f'(u) = \lim_{\Delta u \to 0} \frac{f(u+\Delta u) - f(u)}{\Delta u}, \tag{6}$$

where Δu is an increment in u and $f(u+\Delta u)$ is the corresponding increment in v.

The equation of the point (x_1, y_1) (not on the x-axis) may be written by means of (3)

$$v = -\frac{x_1}{y_1} u - \frac{1}{y_1}.$$

Here the line derivative is a constant $-\frac{x_1}{y_1}$. This is the slope of a line perpendicular to another line having the slope $\frac{y_1}{x_1}$, that is, perpendicular to the radius vector drawn to the point (x_1, y_1). Clearly all points on a line with the origin have the radii vectores of the same slope and consequently have the same line derivative. In point geometry parallel lines had the same slope. These are dual circumstances which we may state thus:

All points collinear with the origin have the same line derivative.	All lines concurrent with the line at infinity have the same point derivative.

The question of interpreting $f'(u)$ for a general curve still remains. Let (u_1, v_1) and $(u_1+\Delta u, v_1+\Delta v)$ be two lines on the curve $v = f(u)$. By (4) the point of intersection of these lines has the equation

$$\begin{vmatrix} u & v & 1 \\ u_1 & v_1 & 1 \\ u_1+\Delta u & v_1+\Delta v & 1 \end{vmatrix} = 0,$$

or

$$\frac{-\Delta v}{\begin{vmatrix} u_1 & v_1 \\ \Delta u & \Delta v \end{vmatrix}} u + \frac{\Delta u}{\begin{vmatrix} u_1 & v_1 \\ \Delta u & \Delta v \end{vmatrix}} v + 1 = 0. \tag{7}$$

The point coordinates of this point are $\left(\frac{-\Delta v}{\begin{vmatrix} u_1 & v_1 \\ \Delta u & \Delta v \end{vmatrix}}, \frac{\Delta u}{\begin{vmatrix} u_1 & v_1 \\ \Delta u & \Delta v \end{vmatrix}} \right)$;

the radius vector drawn to it has the slope $-\frac{\Delta u}{\Delta v}$; the slope of the perpendicular to this radius vector is $\frac{\Delta v}{\Delta u}$. Now as Δu approaches zero,

Δv approaches zero, and the second line approaches the first. The point of intersection of the two lines approaches the point of contact of (u_1, v_1) with the curve $v=f(u)$. The quantity $\frac{\Delta v}{\Delta u}$ approaches, by definition, the line derivative $f'(u_1)$. Therefore, if the variables u and v are connected as in (5), the line derivative "at" the line (u_1, v_1) is the slope of a line perpendicular to the radius vector drawn to the point of contact of (u_1, v_1) and the curve. For simplicity in what follows, the radius vector drawn to the point of contact of (u_1, v_1) and the curve will be called the *contact vector* for the line (u_1, v_1).

The line derivative can be used to find the intercepts of a curve. By setting it equal to zero and solving the resulting equation, one obtains those values of u for which the perpendicular to the contact vector has a slope of zero and is thus parallel to the x-axis. Thus the contact vector must lie on the y-axis. From such a value u_1 one can compute the corresponding v_1; the y-intercept of the curve is then $-\frac{1}{v}$. By changing the equation of the curve to the form $u=\varphi(v)$, setting $\varphi'(v)$ equal to zero, solving for v, computing the corresponding u, and taking its negative reciprocal, one can obtain the x-intercepts of the curve.

Here is manifested a dual relationship between a maximum or minimum point (or a point of inflection where the curve has a horizontal tangent) and a y-intercept. The one is found by setting a point derivative equal to zero; the other is found by treating a line derivative in the same way. A similar relationship exists between vertical tangents and x-intercepts. No longer should the finding of maximum and minimum points be considered the deeper problem merely because calculus is involved. In line geometry they can easily be found (maximum and minimum values of the ordinate, at least) without any resort to calculus.

One can also find the slope of a curve at any point on it. Knowing the point, the slope of the radius vector to it is easily found. The line derivative must now be set equal to the negative reciprocal of that quantity. The resulting equation is solved for u, giving one coordinate of the line; from the equation of the curve the corresponding v is determined, giving the other coordinate. If the coordinates are (u_1, v_1) the slope of the line is $-\frac{u_1}{v_1}$. Obviously, this will not work for finding the slope of a curve at the origin.

In line geometry, however, a more natural problem would be to determine the point of contact of a certain tangent to the curve. This can be done both geometrically and algebraically. Let the curve have the equation

$$v = f(u), \tag{5}$$

and let the line be (u_1, v_1). Obtain the line derivative $f'(u_1)$. This is the negative reciprocal of the slope of the contact vector for (u_1, v_1). Construct, therefore, through the origin a line having the slope $\frac{-1}{f'(u_1)}$. This line will evidently meet the line (u_1, v_1) in the required point of contact.

To obtain the equation of this point of contact, consider the quantity $-\dfrac{1}{f'(u_1)}$. This is the slope of the contact vector for (u_1, v_1) and thus is the ratio of the y-coordinate of the point to the x-coordinate. That means the point has an equation of the form

$$f'(u_1)u - v + k = 0,$$

since

$$\frac{y}{x} = -\frac{1}{f'(u_1)}$$

The point lies on the line (u_1, v_1). Therefore, the proper value of k is that for which the equation

$$f'(u_1)\,u_1 - v_1 + k = 0$$

holds, or

$$k = v_1 - u_1\, f'(u_1).$$

Thus the equation is

$$f'(u_1)\,u - v + v_1 - u_1\, f'(u_1) = 0$$

or

(8) $$\frac{f'(u_1)}{v_1 - u_1 f'(u\ )}u - \frac{1}{v_1 - u_1 f'(u\ )}v + 1 = 0.$$

In this work it is assumed that (u_1, v_1) is not a line through the origin. If k happens to be zero then instead of (8) the equation of the point of contact is

(9) $$f'(u_1)u - v = 0,$$

which represents a family of parallel lines and thus defines the point of contact to be the point at infinity. Obviously, if u_1 and v_1 are finite, k cannot be infinite, for if it were, equation (8) would represent the origin, a point not even on the tangent line. Finally, no real difficulty results if $f'(u_1)$ is either zero or infinite; in the first case the point of contact lies on the y-axis, in the second it lies on the x-axis.

Another form into which (8) can be put is the following one which is reminiscent of the point-slope form of the point equation of a line:

(10) $$v - v_1 = f'(u_1)(u - u_1).$$

Here again the principle of duality is neatly illustrated.

The point equation of the tangent to the curve $y = f(x)$ at the point $(x_1\ \ y_1)$ is

$$y - y_1 = f'(x_1)(x - x_1).$$

The line equation of the point of contact of the curve $v = f(u)$ and the line (u_1, v_1) is

$$v - v_1 = f'(u_1)(u - u_1).$$

An asymptote to a curve may be defined as a tangent to that curve at a point at infinity, provided that tangent is not the line at infinity itself. The line derivative $f'(u)$ "at" a certain line of the curve is the slope of a perpendicular to the contact vector for that line. If the point of contact is on the line at infinity, the contact vector will have the same slope as the asymptote, and the line derivative will be the negative reciprocal of this slope. Thus if the asymptote is the line (u_1, v_1) (having a slope $-\frac{u_1}{v_1}$), the line derivative $f'(u_1)$ will be equal to $\frac{v_1}{u_1}$. Hence to find the asymptotes of a curve, set the line derivative equal to $\frac{v}{u}$ or $\frac{f(u)}{u}$, solve for u, compute the corresponding v's from the equation of the curve, and their coordinates are determined. This method does not apply to asymptotes through the origin since such lines do not have finite coordinates.

If, in considering a curve $y = f(x)$ in point geometry, one set $f'(x)$ equal to $\frac{y}{x}$ or $\frac{f(x)}{x}$, solved for x and computed y, one would obtain the points of the curve at which tangents pass through the origin.

The coordinates of a point on the curve $y = f(x)$ at which the tangent passes through the origin are found by solving the equation

$$f'(x) = \frac{f(x)}{x}$$

for x and computing the corresponding y.

The coordinates of a tangent to the curve $v = f(u)$ whose point of contact lies on the line at infinity are found by solving the equation

$$f'(u) = \frac{f(u)}{u}$$

for u and computing the corresponding v.

If asymptotes to a curve can be found by the above method, then setting $f'(u)$ equal to $-\frac{u}{v}$ and solving with equation (5) for u and v will give the coordinates of a line of the curve not parallel but perpendicular to the contact vector. That particular contact vector, being perpendicular to the tangent, is perpendicular to the curve. Thus the radii vectores which cut the curve orthogonally can be determined (except, of course, where the curve passes through the origin). It is necessary simply to construct through the origin a perpendicular to the tangent obtained by this procedure. If this tangent is (u_1, v_1) the required radius vector will have the slope $\frac{v_1}{u_1}$.

In point geometry, the analogous procedure will yield a point the radius vector through which cuts the curve orthogonally. Thus the problem of determining such radii vectores is very close to being self-dual.

Let the equation of the curve

$$y = f(x)$$

Let the equation of the curve

$$v = f(u)$$

be solved simultaneously with

$$f'(x) = \frac{-x}{f'(x)}$$

to yield, we shall suppose, a solution $x=x_1$, $y=y_1$. These are the coordinates of a point the radius vector through which cuts the curve orthogonally and has the slope $\frac{y_1}{x_1}$.

be solved simultaneously with

$$f'(u) = \frac{-u}{f'(u)}$$

to yield, we shall suppose, a solution $u=u_1$, $v=v_1$. These are the coordinates of a line the contact vector for which cuts the curve orthogonally and has the slope $\frac{v_1}{u_1}$.

The second line derivative $f''(u)$ is defined as the rate of change of $f'(u)$ with respect to u. Consequently, it is the rate of change of the slope of a line perpendicular to the contact vector for a general line of the curve $v=f(u)$. This slope will increase, decrease, or remain the same as u increases, according as $f''(u)$ is positive, negative, or zero. But the slope of the contact vector itself will likewise increase, decrease, or remain the same under those circumstances. Thus, for a finite value u_1 of u such that $f''(u_1)=0$, the slope of the contact vector has a stationary value. That is, as u ranges over an interval of values containing u_1, the contact vector rotates about the origin in one direction, stops, and rotates back in the opposite direction, the stopping place corresponding to $u=u_1$. In other words, the contact vector for the line whose u-coordinate is u_1 (and we shall call its v-coordinate v_1) has a maximum or a minimum inclination to the x-axis. (This may be only a relative maximum or minimum.) Thus at least a limited part of the curve in the neighborhood of $u=u_1$ lies entirely on one side of this contact vector. The curve cannot be tangent to this contact vector (even though it would then lie on one side), for then u_1 and v_1 would both have to be infinite (or one of them infinite and the other indeterminate). Thus at the point of contact of the line (u_1, v_1) with the curve there must be a cusp. It is easy to see how a cusp could arise quite naturally in a curve which is generated by a straight line. There must be no corner (point at which a finite angle is made by two tangents to the curve) at this point, for v must be a continuous function of u. Moreover, the cusp must be such that the tangent (u_1, v_1) lies between its two branches as shown in figure 1 and not as in figure 2, for in the latter case v would not be a single-valued function of u in the neighborhood of $u=u_1$. This is easily seen by the figure.

Then by setting $f''(u)$ equal to zero, solving for u, and computing the corresponding v, one can obtain the coordinates (u_1, v_1) of a tangent to the curve through a cusp on it. By methods already noted, the point of contact of this tangent with the curve can be found. This will be the vertex of the cusp. It is possible to obtain more information about this cusp — that is, to determine on which side of the

contact vector it lies. If, as u increases through u_1, the contact vector rotates first counter-clockwise and then clockwise, $f''(u)$ must be first positive, then zero, and then negative. That is, it decreases, and the rate of change of $f''(u)$ with respect to u is negative at $u=u_1$.

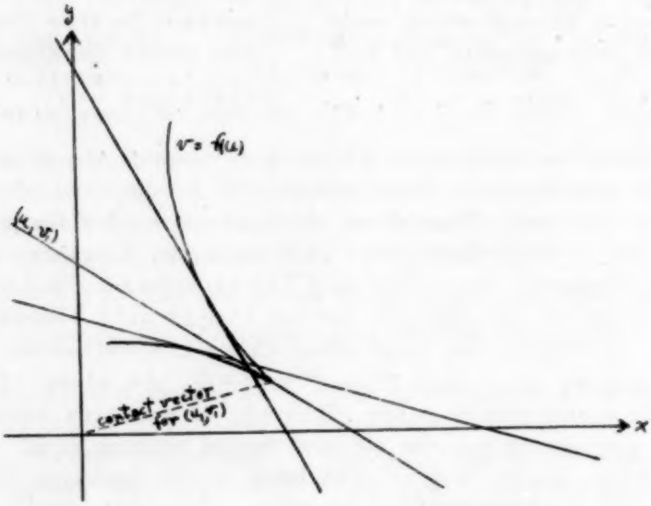

Fig. 1. One value of v corresponds to one value of u.

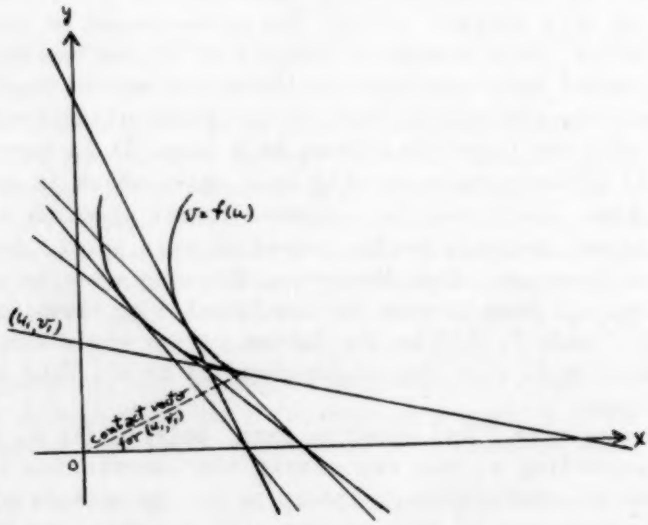

Fig. 2. Two values of v correspond to one value of u.

But this rate of change is precisely $f'''(u_1)$. Thus if $f'''(u_1)$ is negative, the cusp lies on the clockwise side of the contact vector for (u_1, v_1) (on the right side as it would appear to an observer standing at the origin and looking towards the vertex of the cusp). A similar argument will show that the cusp lies on the other side when $f'''(u_1)$ is positive. In case $f'''(u_1) = 0$, the test fails.

The coordinates of a point of inflection on the curve $y = f(x)$ are found by solving $f''(x) = 0$ for x and computing the corresponding y. The slope of the curve at this point is found by methods already developed. If the x-coordinate of this point is x_1, the curve is rising or falling through this point according as $f'''(x_1)$ is positive or negative.

The coordinates of a tangent through a cusp on the curve $v = f(u)$ are found by solving $f''(u) = 0$ for u and computing the corresponding v. The vertex of this cusp (point of contact of this tangent) is found by methods already developed. If the u-coordinate of this tangent is u_1, the cusp lies on the counter-clockwise or clockwise side of the contact vector for this tangent according as $f'''(u_1)$ is positive or negative.

It is almost intuitively evident that a point of inflection and a cusp are duals of each other if the curves containing them are thought of as being generated by a straight line and a point, respectively. Both the line and the point move in one manner, stop, and then move in the opposite manner. The line rotates first in one direction and then in the other; the point moves first in one direction and then in the other.

In this paper only plane geometry and functions of a single variable have been considered. There is, however, a principle of duality in three-dimensional space in which the point and the plane are duals of each other. An analytic geometry can here be developed, based upon the plane as the fundamental element and considering surfaces, curves, and points as made up of planes. Coordinates commonly used to represent a plane are the negative reciprocals of its intercepts on the x-, y-, and z-axes (designated by the letters u, v, w). As before, elements (in this case planes) through the origin are not considered. In this type of geometry calculus also finds a place, and interpretations and applications of the partial planar derivatives, $\frac{\partial w}{\partial u}$ and $\frac{\partial w}{\partial v}$, can be devised.

References

1. W. C. Graustein, *Introduction to Higher Geometry*, Macmillan Co., 1937, pp. 26-7, 65, 70-1, 202-7, 404-8.

2. A. Clement-Jones, *An Introduction to Algebraical Geometry*, Oxford, Clarendon Press, 1912, chapter 10.

3. G. Salmon, *A Treatise on Conic Sections*, Longmans, Green and Co., tenth ed., 1896, p. 278.

Beloit College, Beloit, Wisconsin.

TEACHING OF MATHEMATICS

Edited by

Joseph Seidlin and C. N. Shuster

This department is devoted to the teaching of mathematics. Thus articles on methodology, exposition, curriculum, tests and measurements, and any other topic related to teaching, are invited. Papers on any subject in which you, *as a teacher*, are interested, or questions which you would like others to discuss, should be sent to Joseph Seidlin, Alfred University, Alfred, New York.

ADVANCED MATHEMATICS FROM AN ELEMENTARY STANDPOINT

M. R. Spiegel

There are many results in the field of mathematics which, however beautiful and inspiring they may be, have never been revealed to the student who does not go beyond a first course in the calculus. The author has found that many of these results can be obtained with but a few mathematical tools and a little imagination. The methods, it cannot be denied, are somewhat heuristic, but the student can be made aware that he is essentially "discovering mathematics" (with, of course, the aid of a teacher) and that a "revarnishing with rigour" is to take place after an intuitive understanding of the problem has been achieved in order to make a derivation complete in all details.

Two interesting results, formerly falling into the category of "advanced techniques" are the following well-known (to the advanced) series

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}$$

$$\frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \frac{\pi^4}{90}$$

Normally, to derive these, one needs at least a background of Fourier series and, if one is somewhat ingenious, one can derive them with the aid of integral equations or complex variable theory.

It is the purpose of this article to show how, not only these results, but other very interesting and useful results may be derived with only the aid of simple elementary algebra and elementary calculus.

The student is aware, first of all, that if a polynomial vanishes for $x = a$ then the polynomial has $(x - a)$ as a factor. A little imagination plus some facility for generalization will make plausible the following question. If the theorem is true for polynomials, why could it not be true for any function — for example $\sin x$ (to take the next complicated function after polynomials)? If the theorem is true, then it should be possible, since $x = 0, \pm\pi, \pm 2\pi, \cdots$ are the values of x which make the function $\sin x$ vanish, to write

$$\sin x \stackrel{?}{=} x(x^2 - \pi^2)\,(x^2 - (2\pi)^2)(x^2 - (3\pi)^2)\cdots \tag{1}$$

(Here the question mark surmounts the equals sign always as a symbol of doubt since we do not know yet whether the conjecture is correct.)

We now adopt the simple expediency of substituting a value for x to see if "both sides of (1) are equal". In a sense we are "performing an experiment". A simple value to substitute is $x = \pi/6$ (30° expressed in radians). Performing the experiment, we see that both sides are not equal. Nevertheless, courage and a faith in the validity of our conjecture prevent us from resignation (plus, of course, the faith in the teacher). Suddenly the inspiration of rearranging the right hand side of (1) presents itself and after some more experimentation we finally arrive at the possibility

$$\sin x \stackrel{?}{=} x\left[1 - \frac{x^2}{\pi^2}\right]\left[1 - \frac{x^2}{(2\pi)^2}\right]\cdots$$

Substitution of $x = \pi/6$ on both sides and a subsequent comparison makes us convinced that the formula is correct but, as elementary investigators, with none of the powerful tools of analysis to give us support, we cannot yet remove the symbol of interrogation from over the equals sign. The more values of x that we try, however, the more we are convinced of the validity of our newly derived formula. When we are convinced enough (just as the physicist is when his set of experiments has yielded some unification of results) we decide to remove the question mark from the paper. As mathematicians, however, we keep it only in our minds and write

$$\frac{\sin x}{x} = \left[1 - \frac{x^2}{\pi^2}\right]\left[1 - \frac{x^2}{(2\pi)^2}\right]\cdots \tag{2}$$

where we have divided by x only because it seems so different from the other factors on the right hand side. If we multiply all of the factors on the right hand side of (2) in the manner we were taught in elementary algebra we obtain

$$\begin{aligned}\frac{\sin x}{x} &= 1 - \frac{x^2}{\pi^2}\left[\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\right] \\ &\quad + \frac{x^4}{\pi^4}\left[\frac{1}{1^2 2^2} + \frac{1}{1^2 3^2} + \cdots + \frac{1}{2^2 3^2} + \cdots\right] - \cdots\end{aligned} \tag{3}$$

Now we realize that $\sin x$ can be written

$$\sin x = a + bx + cx^2 + dx^3 + \cdots$$

where the constants $a, b, c, \cdots$ may be determined in the usual way using Maclaurin series as in elementary calculus. Making use of this we find

$$\text{(3)} \qquad \frac{\sin x}{x} = 1 - \frac{x^2}{6} + \frac{x^4}{120} - \cdots$$

Comparing (2) with (3) we seem to have no choice but to observe that

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6} \equiv S_1$$

(our first formula) and

$$\frac{1}{1^2 2^2} + \frac{1}{1^2 3^2} + \cdots + \frac{1}{2^2 3^2} + \frac{1}{2^2 4^2} + \cdots = \frac{\pi^4}{120} \equiv S_2$$

which is not our second result.

We do, however, notice that

$$S_1^2 - 2S_2 = \frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots$$

as can be seen by actually performing the indicated operations. It follows therefore that

$$\frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \frac{\pi^4}{36} - \frac{2\pi^4}{120} = \frac{\pi^4}{90}$$

(our second result). It is possible also to extend the method to evaluate the series $\frac{1}{1^6} + \frac{1}{2^6} + \cdots$ and other series with even exponent. It will be observed that we have not, by this method, obtained a closed form for the series

$$\frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots$$

In fact even the most advanced techniques have failed in this respect.

The precise conditions, or if you like, reasons why the above method works, are available to the mathematician. He need only go to any advanced textbook to find out. Nevertheless, it is surprising how simple the method is and how useful it is for obtaining many results. For example we may find by the method the sum of reciprocal even powers of the zeros of the following functions (a) $J_0(x)$ (b) $J_n(x)$, $n>0$ (c) $\cosh x \cos x + 1$ (d) $\tan x - x$ to mention just a few. Part (c) comes up in the theory of vibrations of an elastic bar as investigated by Lord Kelvin.

To summarize, I would like to suggest that just as Felix Klein has asked us to reexamine "Elementary Mathematics From An Advanced Standpoint", it might be a good idea to present to our students some "Advanced Mathematics from an Elementary Standpoint".

Rensselaer Polytechnic Institute.

A NOTE ON THE USE OF DISCRIMINANTS

Robert E. Horton

Many beginning students of mathematics make the mistake of learning techniques without gaining an understanding of the reasons underlying those techniques. An illustration follows which shows how the unwary student might be trapped by a blind dependence upon the discriminant of a quadratic.

Suppose the student is required to find the members of the family of parabolas $kx - y^2 = 9$ which are tangent to the circle $x^2 + y^2 = 16$. The usual first reaction might be to solve the pair simultaneously getting a quadratic in x containing k in one or more of the coefficients. Then setting the discriminant of this quadratic equal to zero he might expect to obtain the conditions on k which would yield the required parabola. Proceeding to do this we have

$$\begin{array}{r} x^2 + y^2 = 16 \\ kx - y^2 = 9 \\ \hline x^2 + kx = 25 \end{array}$$

or $x^2 + kx - 25 = 0$. Then set the discriminant of this equation equal to zero; that is, $k^2 + 100 = 0$. This equation has as roots $k = \pm 10i$.

Here the beginning student might draw the conclusion that there is no member of the family of parabolas with real values of k which will be tangent to the circle. However, the student upon sketching the circle and several members of the family of parabolas would see that a parabola with real k certainly should be expected from the general configuration.

Somewhat disturbed but still hopeful of finding the real k he proceeds to eliminate x from the two equations. Doing so a quartic in y is obtained of the form $y^4 + (18 + k^2)y^2 + (81 - 16k^2) = 0$. If the student considered this as a quadratic in y^2 and set the resulting discriminant equal to zero he would have $(18 + k^2)^2 - 4(81 - 16k^2) = 0$ or $k^2(k^2 + 100) = 0$. Thus he again obtains as roots $k = \pm 10i$ and $k = 0$.

With this evidence he is convinced that there is no real k which will make for tangency as $k = 0$ obviously will not work.

However, by trial and error he might find out that for $k = \pm 9/4$ he has obtained real members of the family of parabolas which are tangent to the circle. Our unwary student now is about ready to assume that the method of equating discriminants to zero is an unreliable one.

It is at this point that the instructor must be careful to point out what is wrong with the student's application of a technique to a situation where it can not be expected to provide results. In the first use of the discriminant of the quadratic in x, the student should have noted that due to symmetry the two x coordinates of the intersections of the circle and *any* member of the parabola family would be equal. Thus the discriminant will not yield any information.

In the second application of the discriminant the student again should observe from symmetry that the coordinates of intersection of any member of the parabola family with the circle would be of the from (x_0, y_0) and $(x_0, -y_0)$. Now any quadratic in y with roots $\pm y_0$ must be of the type $y^2 - p = 0$ and the discriminant of this can not be zero unless p itself is zero.

It is this last possibility which proves to be the case. For if the equation $y^4 + (18 + k^2)y^2 + (81 - 16k^2) = 0$ is solved for y^2 and then for y we have $y = \pm\sqrt{\frac{1}{2}(-18 - k^2 \pm k\sqrt{k^2 + 100})}$. For these roots to be equal they must be zero. Thus the conditions on k are

$$-(18 + k^2) \pm k\sqrt{k^2 + 100} = 0$$

The roots of this equation are $k = \pm 9/4$. These are the values of k for which the student has been searching.

A careful observer would have seen from symmetry the necessity for the point of tangency to have a zero y coordinate and thereby could have obtained the result $k = \pm 9/4$ almost immediately.

Los Angeles City College.

AN ELEMENTARY INTRODUCTION TO THE CALCULUS OF VARIATIONS

Magnus R. Hestenes

1. *Introduction.* Calculus of variations is concerned primarily with the study of maxima and minima of real valued functions. In the elementary calculus we study the problem of maximizing or minimizing a function $f(x)$ of a single variable x. The results are then extended to a function $f(x,y)$ of two variables x and y. Next functions $f(x,y,z)$ of three variables are considered, and so on. In the calculus of variations the variables need not be points but may be arcs or surfaces. For example, we may wish to find an arc of shortest length joining two given points or a surface of least area having a given boundary. Because the variables are in general arcs and surfaces, it is to be expected that the technical details may become rather involved. However, the problems considered are easily described and the main results easily understood. The present paper will be devoted to description of typical problems together with some of the main results. We shall restrict ourselves to the case when the independent variables are curves.

Perhaps the simplest problem in the calculus of variations is that of finding arcs of least length satisfying prescribed conditions. The curve of least length joining two given points is of course a straight line segment. If the curves are restricted to lie on a sphere, the curve of least length is an arc of a great circle. What the curves of least length (called geodesics) are on a torus or on a more complicated surface is not at all obvious and their study forms an interesting chapter in the calculus of variations and differential geometry.

It is an interesting fact that the theory of the calculus of variations did not arise initially through the study of curves of least length. It arose instead through the study of paths of least time, called Brachistochrones. In 1696 the Bernoulli brothers studied the following problem: Consider a bead of unit mass sliding along a wire joining two fixed points and subject only to gravitational force. How shall the wire be shaped so that the bead will traverse the wire in the least time? It can be shown that the wire must take the shape of an inverted cycloid. Through the study of this and related problems the Bernoulli brothers followed by Euler and Lagrange began the development of the modern theory of the calculus of variations.

In the following pages we shall in general restrict ourselves to minimizing arcs. The corresponding results for maximizing arcs are obtained by reversing inequalities whenever they occur.

Typical questions that arise in the calculus of variations are the following:

1. Does the minimum problem have a solution?

2. If it has a solution what properties does the solution have and how can the solution be found?

3. What properties will insure that a given arc be a solution of a given problem?

4. Does a problem have various types of solutions and how are they to be classified?

We do not propose to answer all of these questions here but shall content ourselves with remarks that are pertinent relative to the second question.

2. *Problems in the plane.* Consider now an arc C joining two points 1 and 2 as shown in the accompanying figure. There are infinitely many

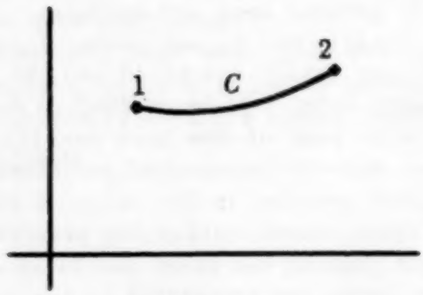

Fig. 2.1

arcs of this type. The length of C is given by the formula

$$I(C) = \int_{x_1}^{x_2} (1 + y'^2)^{1/2} dx. \tag{2:1}$$

Here and elsewhere y' denotes the slope dy/dx of the curve at the point (x,y). Thus, the arc of shortest length is one which minimizes $I(C)$. If we ask for the arc joining 1 and 2, which when revolved about the x-axis generates the surface of revolution of minimum area, then this arc must minimize the integral

$$I(C) = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx. \tag{2:2}$$

This integral when multiplied by 2π represents the area of the surface of revolution generated by C.

Let us now consider a path of least time joining two given points 1 and 2 subject only to gravitational force. To this end it is convenient to choose the y-axis to be the vertical axis directed downward as indicated in Figure 2:2. Let us suppose therefore that we have a bead P sliding along C from 1 to 2 with an initial speed v_1 at 1. If we

denote by s the distance traversed in t seconds and by v the speed

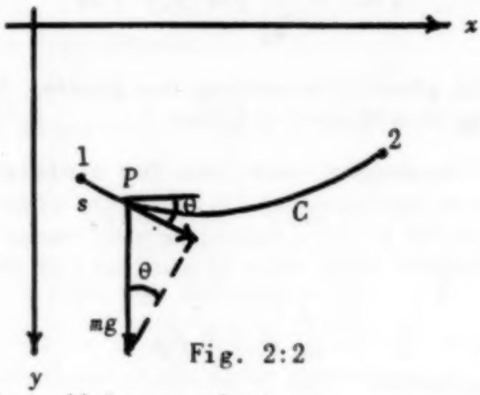

Fig. 2:2

ds/dt, we have the well known relations

$$m \frac{dv}{dt} = mg \sin\theta = mg \frac{dy}{ds} \quad \text{or} \quad v\,dv = g\,dy.$$

Here m is the mass of the bead and g is the gravitational acceleration. Consequently, by integration, we have

$$v^2 = 2gy + b.$$

Since $y = y_1$, $v = v_1$ at $t = 0$

$$v^2 = 2g(y - y_0), \quad y_0 = y_1 - \frac{v_1^2}{2g}.$$

Let us select the x-axis so that $y_0 = 0$. Then the time T elapsed in traversing C from 1 to 2 is given by the integral

$$T = \int_0^l \frac{ds}{v} = \frac{1}{\sqrt{2g}} \int_{x_1}^{x_2} \left[\frac{1 + y'^2}{y}\right]^{1/2} dx.$$

The path of least time is accordingly obtained by minimizing the integral

$$I(C) = \int_{x_1}^{x_2} y^{-1/2} (1 + y'^2)^{1/2}\,dx. \tag{2:3}$$

in the class of arcs joining 1 and 2.

The form of the integrals (2:1), (2:2) and (2:3) suggest that they can be studied simultaneously by studying the integral

$$I(C) = \int_{x_1}^{x_2} y^r (1 + y'^2)^{1/2}\,dx \qquad (y > 0) \tag{2:4}$$

or even more generally by studying the integral

$$I(C) = \int_{x_1}^{x_2} f(x,y,y')\,dx \tag{2:5}$$

where $f(x,y,y')$ is given. Concerning the general integral (2:5) we have the following result due to Euler:

I. **The first necessary condition for a minimum.** *Let C be an arc which minimizes the integral* (2:5) *in the class of arcs joining two given points* 1 *and* 2. *If C possesses continuous first and second derivatives with respect to x, then it must satisfy the Euler equations*

$$\frac{d}{dx} f_{y'} = f_y \tag{2:6}$$

and in fact the equation

$$\frac{d}{dx}(f - y' f_{y'}) = f_x, \tag{2:7}$$

which is equivalent to (2:6) *whenever $y' \neq 0$. A solution of equation* (2:6) *is called an extremal.*

The proof of this result will be given in the next section. Incidentally we have used the abbreviations

$$f_x = \frac{\partial f}{\partial x}, \quad f_y = \frac{\partial f}{\partial y}, \quad f_{y'} = \frac{\partial f}{\partial y'}\,.$$

Let us apply this result to the integral (2:4) in which

$$f = y^r(1+y'^2)^{1/2}.$$

Here

$$f_y = r\, y^{r-1}(1+y'^2)^{1/2}, \qquad f_{y'} = y^r\, y'(1+y'^2)^{-1/2},$$

$$f_x = 0, \qquad f - y' f_{y'} = y^r(1+y'^2)^{-1/2}.$$

The equations (2:6) and (2:7) are respectively equivalent to the equations

$$y'' = r\,\frac{1+y'^2}{y} \tag{2:8}$$

$$y^r(1+y'^2)^{-1/2} = b^r \tag{2:9}$$

where b is a constant. We restrict ourselves to the case when $y > 0$. Then equation (2:8) tells us that along the minimizing arc y'' has the same sign as r. Thus the arc is concave upward when $r > 0$ and concave downward when $r < 0$. The extremals are straight lines when $r = 0$.

In order to integrate (2:8) we use its integral (2:9) giving us

$$y = b(1+y'^2)^{\frac{1}{2r}}.$$

Letting $y' = \tan\theta$ we obtain

$$y = b(\sec\theta)^{\frac{1}{r}}, \quad dy = \frac{b}{r}(\sec\theta)^{\frac{1}{r}}\tan\theta\, d\theta.$$

Consequently

$$dx = \frac{dy}{y'} = \frac{b}{r}(\sec\theta)^{\frac{1}{r}}\, d\theta.$$

An integration yields the following:

If the integral to be minimized is of the form (2:4) with $r \neq 0$, the extremals (solutions of the Euler equations) are given parametrically by the equations

$$(2{:}10) \qquad x = a + \frac{b}{r}\int_0^\theta (\sec t)^{\frac{1}{r}}\, dt, \quad y = b(\sec\theta)^{\frac{1}{r}},$$

where θ is the angle of inclination. An alternate form of these equations is

$$(2{:}11) \qquad x = a + \frac{b}{r}\int_0^\phi (\cosh u)^{\frac{1-r}{r}}\, du, \quad y = b(\cosh\phi)^{\frac{1}{r}}.$$

The equations (2:1) are obtained by making the substitution $y' = \sinh\phi$ in place of $y' = \tan\phi$.

The character of the extremal family (2:10) depends in an essential way on the sign of r.

If $r < 0$, the extremals (2:10) are concave downward, the line $x = a$ being a line of symmetry. Through every pair of points 1 and 2 in the upper half plane there passes one and but one extremal (2:10). This extremal minimizes the integral (2:4) in the class of arcs C joining 1 and 2 and lying in the upper half plane. If $r = -1$, the extremal (2:10) is a circular arc

$$(2{:}12) \qquad x = a - b\sin\theta, \quad y = b\cos\theta$$

of radius b with $(a,0)$ as its center. If $r = -\frac{1}{2}$ the extremal (2:10) is an arc of the cycloid

$$(2{:}13) \qquad x = a - \frac{b}{2}(2\theta + \sin 2\theta), \quad y = \frac{b}{2}(1 + \cos 2\theta)$$

generated by a point P on a circle of diameter b rolling on the x-axis. This cycloid is accordingly a path of least time.

The equations (2:12) and (2:13) are obtained from (2:10) by setting $r = -1$, $r = -\frac{1}{2}$ respectively and integrating. It is clear that through

every pair of points in the upper half plane there is but one circular arc having its center on the x-axis. Consequently, in the case $r = -1$ there is a unique extremal joining a pair of points in the upper half plane. This result remains true whenever $r < 0$ but it is more difficult to prove and we shall not do so here. If we accept the fact that minimizing arcs exist when $r < 0$, then the extremals must be minimizing since they are the only ones that satisfy the first necessary condition.

The situation is quite different when $r > 0$. A complete statement of the result can be made as follows.

If $r > 0$, the extremals (2:10) are concave upward and the line $x = a$ is a line of symmetry. Given a pair of points 1 and 2 in the upper half plane there may exist two, one or no extremals joining them. If two extremals exist, then one is minimizing and the other is not. If but one extremal exists, it fails to be minimizing. If no extremal joins 1 and 2, there is no minimizing arc above the x-axis. If $r = \frac{1}{2}$ the extremals (2:10) are parabolas

$$y = b[1 + \tfrac{1}{4}\left[\frac{x-a}{b}\right]^2] . \tag{2:14}$$

If $r = 1$, the extremals (2:11) are catenaries

$$y = b \cosh \left[\frac{x-a}{b}\right] \tag{2:15}$$

and are the extremals for the problem of finding the arc generating a surface of revolution of minimum area.

The general situation is like that of the particular case $r = \frac{1}{2}$. Consequently we shall restrict ourselves to this case. The parabolas (2:14) are obtained from (2:10) with $r = \frac{1}{2}$ by integration. They have the x-axis as the directrix. If we select

$$b = \cos^2\alpha, \quad a = -2 \sin\alpha \cos\alpha$$

equation (2:14) takes the simple form

$$y = 1 + x \tan\alpha + x^2 \frac{\sec^2\alpha}{4} . \tag{2:16}$$

This represents the totality of extremals through the point (1,0). This family has the parabola

$$y = \frac{x^2}{4}$$

as an envelope as shown in figure 2:3. To verify this we differentiate with respect to α and obtain

$$0 = x \sec^2\alpha + \frac{x^2}{2} \sec^2\alpha \tan\alpha \quad \text{or} \quad \tan\alpha = -\frac{2}{x}.$$

Using this result in (2:16) we see that the envelope is given by $y = \dfrac{x^2}{4}$.

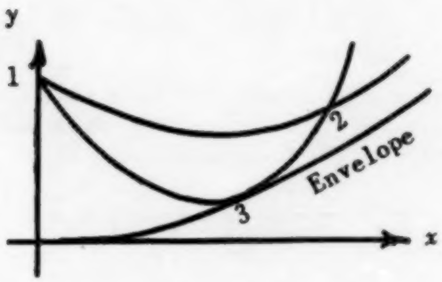

Fig. 2:3

If the point 2 is located to the right of the envelope there is no extremal joining 1 and 2. If it is on the envelope there is but one extremal joining 1 to 2. If it is to the left of the envelope there are two extremals joining 1 and 2 as indicated in Fig. 2:3. The segment that fails to have a point of contact with the envelope is minimizing and the other is not. We are not in position here to establish the minimizing properties of these extremals.

In this section we have brought out the following facts.

1. In order that an arc be a minimizing arc or a maximizing arc it must satisfy the Euler equations (unless it is on the boundary of the domain of validity of our problem).

2. A solution of the Euler equation need not be a minimizing or a maximizing arc.

3. Problems exist for which there is no solution in the sense that there is no extremal satisfying the conditions of the problem.

4. Further criteria must be obtained for distinguishing maximizing and minimizing arcs.

3. *Derivation of the Euler equations.* Some of the techniques used in the calculus of variations are illustrated by the following derivation of Euler's equation. To this end let

$$C: \qquad y(x) \qquad x_1 \leqq x \leqq x_2$$

be an arc that minimizes the integral

$$I = \int_{x_1}^{x_2} f(x,y,y')\,dx$$

in the class of arcs joining its end points. If $\eta(x)$ vanishes at x_1 and x_2, the arc

$$y(x) + b\,\eta(x) \qquad x_1 \leqq x \leqq x_2$$

has the same end points as C for all values of b and contains C for $b=0$. The value of the integral I along this arc is

$$I(b) = \int_{x_1}^{x_2} f(x, y + b\eta,\ y' + b\eta')\,dx.$$

Since $I(0)$ is the value of I along C we have $I(b) \geqq I(0)$, by virtue of the minimizing property of C. Consequently, from calculus we know that the derivative of $I(b)$ at $b = 0$ is zero, that is,

$$I'(0) = \int_{x_1}^{x_2} \{f_y \eta + f_{y'} \eta'\}\,dx = 0\,. \tag{3:1}$$

The expression

$$I_1(\eta) = \int_{x_1}^{x_2} \{f_y \eta + f_{y'} \eta'\}\,dx$$

with $f_y, f_{y'}$ evaluated along C is called the *first variation* of I along C. By virtue of the relation

$$\frac{d}{dx}(f_{y'}\,\eta) = f_{y'}\,\eta' + \eta\,\frac{d}{dx} f_{y'}$$

we have

$$I_1(\eta) = \int_{x_1}^{x_2} \eta\Big(f_y - \frac{d}{dx} f_{y'}\Big)dx + f_{y'}\,\eta\,\Big|_{x_1}^{x_2}\,.$$

Since $\eta(x_1) = \eta(x_2) = 0$, it follows from (3:1) that

$$I_1(\eta) = \int_{x_1}^{x_2} \eta\Big(f_y - \frac{d}{dx} f_{y'}\Big)dx = 0$$

for all functions $\eta(x)$ with continuous derivatives that vanish at x_1 and x_2. This is possible only in case

$$f_y = \frac{d}{dx} f_{y'}\,,$$

as was to be proved. The equation (2:7) follows from the identity

$$\frac{d}{dx}(f - y' f_{y'}) - f_x + y'\Big[\frac{d}{dx} f_{y'} - f_y\Big] = 0\,.$$

The function $\eta(x)$ used in this derivation is called a *variation* and is frequently denoted by δy. If we denote by δI the first variation of I, then the formula for the first variation suggests the formalism

$$\delta I = \delta \int_{x_1}^{x_2} f\ dx = \int_{x_1}^{x_2} \delta f\ dx = \int_{x_1}^{x_2} \{f_y \delta y + f_{y'} \delta y'\}\,dx$$

which remains valid here as long as δy is restricted to vanish at x_1 and x_2. This formalism is very useful but care must be exercised to use it properly. For example the form of δI and the properties of δy are altered if the problem is changed.

Incidentally, since $I(b) \geqq I(0)$ we not only have $I'(0) = 0$ but also $I''(0) \geqq 0$, where $I''(b)$ is the second derivative of $I(b)$. Computing $I''(0)$ and calling the result $I_2(\eta)$ we obtain

$$(3:2) \qquad I_2(\eta) = \int_{x_1}^{x_2} \{f_{yy}\eta^2 + 2f_{yy'}\eta\eta' + f_{y'y'}\eta'^2\}dx \geqq 0,$$

where the derivatives of f are evaluated along C. This integral is called the *second variation* of I along C. We have accordingly the following further necessary condition.

If C is a minimizing arc the inequality (3:2) *must hold along C for every variation $\eta(x)$ having $\eta(x_1) = \eta(x_2) = 0$. If C is a maximizing arc, the inequality in* (3:2) *must be reversed.*

4. *Generalizations to n-dimensions.* The results discribed in the preceding sections can be generalized at once to the problem of minimizing an integral

$$I(C) = \int_{x_1}^{x_2} f(x, y_1, \cdots, y_n, y_1', \cdots, y_n')\, dx$$

in a class of arcs

$$y_i(x) \qquad (x_1 \leqq x \leqq x_2;\ i = 1, \cdots, n)$$

lying in $(n+1)$-dimensional space of points $(x, y_1, \cdots, y_n)$. By varying each component $y_i(x)$ separately it is seen that the Euler equations corresponding to (2:6) take the form

$$(4:1) \qquad \frac{d}{dx} f_{y_i'} = f_{y_i} \qquad (i = 1, \cdots, n).$$

Every solution of these equations also satisfies the equation

$$(4:2) \qquad \frac{d}{dx}(f - y_i' f_{y_i'}) = f_x,$$

in which the repeated index denotes summation with respect to that index. This follows from the identity

$$\frac{d}{dx}(f - y_i' f_{y_i'}) - f_x + y_i' \left(\frac{d}{dx} f_{y_i'} - f_{y_i}\right) = 0$$

which can be verified by carrying out the indicated differentiations. As before, a maximizing or a minimizing arc must satisfy Euler's equations (4:1) and (4:2).

In three-dimensional space with coordinates (x, y, z) the integral $I(C)$ is of the form

$$I(C) = \int_{x_1}^{x_2} f(x, y, z, y', z')\, dx$$

and the Euler equations (4:1) and (4:2) become

$$
\begin{aligned}
&\frac{d}{dx} f_{y'} = f_y, \qquad f_{z'} = f_z \\
(4{:}3) \qquad &\frac{d}{dx}(f - y'f_{y'} - z'f_{z'}) = f_x .
\end{aligned}
$$

One of the important applications of the results described above lies in the field of mechanics. Let $U(x,y,z)$ be a potential function associated with a conservative field of force. Then the equations

$$
(4{:}4) \qquad m x'' = -U_x, \quad m y'' = -U_y, \quad m z'' = -U_z ,
$$

where primes denote time derivatives, are the differential equation of motion for a particle of mass m moving in this field. We have the following important principle:

Hamilton's principle. *The differential equations* (4:4) *are the Euler equations of the integral*

$$
I = \int (T-U)\,dt
$$

where T is the kinetic energy

$$
T = \frac{m}{2}(x'^2 + y'^2 + z'^2)
$$

of the moving particle described above.

The function $L = T-U$ is called the Lagrangian. An easy computation will show that the equations

$$
\frac{d}{dt} L_{x'} = L_x, \quad \frac{d}{dt} L_{y'} = L_y, \quad \frac{d}{dt} L_{z'} = L_z
$$

are identical with the equations (4:4). Moreover since L is independent of t, the analogue of equation (4:2) becomes

$$
\frac{d}{dt}(L - x'L_{x'} - y'L_{y'} - z'L_{z'}) = 0 .
$$

Hence

$$
L - x'L_{x'} - y'L_{y'} - z'L_{z'} = -(T+U) = \text{constant}
$$

along the trajectory of the particle. Thus, we have established anew the significant result that along the trajectory of a particle moving in a conservative field of force the *sum of the kinetic and potential energy is a constant.*

One of the advantages of Hamilton's Principle is that it is independent of the coordinate system used. As an illustration consider a particle acted upon by a central force inversely proportional to the square of the distance from the center. Assuming planar motion and using polar coordinates with the pole at the center of force we have

$$T = \frac{c}{r}\,(\dot{r}^2 + r^2\,\dot{\theta}^2), \quad U = -\frac{c}{r}$$

Letting $L = T - U$ the Euler equations become

$$\frac{d}{dt} L_{\dot{r}} = L_r$$

$$\frac{d}{dt} L_{\dot{\theta}} = L_\theta$$

and

$$T + U = h = \text{constant}\,.$$

Since L is independent of θ, the second equation yields

$$L_{\dot{\theta}} = m\,r^2\,\dot{\theta} = \text{constant} \quad \text{or} \quad r^2\,\dot{\theta} = b\,.$$

Inasmuch as $\dot{r} = \dfrac{dr}{d\theta}\,\dot{\theta}$ we see that

$$T = \frac{m}{2}\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right]\dot{\theta}^2 = \frac{m}{2}\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right]\frac{b^2}{r^4}\,.$$

Assuming that $b \neq 0$ the relation $T + U = h$ yields

$$a^2\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right] = h\,r^4 + c\,r^3, \quad a^2 = \frac{mb^2}{2}$$

which when integrated has a solution of the form

$$r = \frac{k}{1 - e\,\cos\,(\theta - \gamma)}\,,$$

where k, e and γ are constants. The path is accordingly a conic.

5. *Parametric problems.* In Section 2 we considered only arcs for which the y coordinate can be expressed as a single real valued function $y(x)$ of the x-coordinate x. These are commonly called non-parametric arcs. In most of the examples however we expressed the minimizing arc in parametric form. Let us now admit as comparison arcs those which coordinates x and y are expressible in terms of a parameter t. Thus the system

$$C \qquad x(t), \quad y(t) \qquad (t_1 \leqq t \leqq t_2)$$

defines a parametric arc provided $x(t)$ and $y(t)$ are continuous and have piecewise continuous derivatives x' and y' with $x'^2 + y'^2 \neq 0$. The integral to be minimized is of the form

$$I(C) = \int_{t_1}^{t_2} f(x, y, x', y')\,dt\,,$$

where f has the property that

$$f(x,y,kx',ky') = k\,f(x,y,x',y') \qquad (\text{all } k>0).$$

This condition implies that $I(C)$ is independent of the parametric representation. For the Brachistochrone problem described in Section 2 we have

$$I(C) = \int_{t_1}^{t_2} y^{-\frac{1}{2}}\,(x'^2 + y'^2)^{\frac{1}{2}}\,dt\,.$$

The corresponding n-dimensional problem is the following one: *To find in the class of parametric arcs*

$$C \qquad y_i(t) \qquad (t_1 \leq t \leq t_2;\ i \quad 1, \cdots, n)$$

joining two fixed points 1 *and* 2, *to find one which minimizes an integral*

$$I(C) = \int_{t_1}^{t_2} f(y,y')\,dt = \int_{t_1}^{t_2} f(y_1, \cdots, y_n, y_1', \cdots, y_n')\,dt\,,$$

where

$$f(y,ky') = k\,f(y,y') \qquad (\text{all } k>0). \tag{5:1}$$

The Euler equations take the form

$$\frac{d}{dt} f_{y_i'} = f_{y_i} \qquad (i = 1, \cdots, n). \tag{5:2}$$

Every minimizing arc must satisfy the Euler equations.

Parametric problems are particularly useful in the study of geodesics on surfaces and of the principle of least action in mechanics. However we shall not pause to discuss them here but will content ourselves with the application to be made in the next section.

6. *Isoperimetric problems.* One of the oldest problems in the calculus of variations is that of finding the shape of a simple closed curve of given length which encloses the largest area. Analytically this problem is that of maximizing the area integral

$$I(C) = \tfrac{1}{2} \int_{t_1}^{t_2} (xy' - yx')\,dt$$

in the class of simply closed curves

$$C: \qquad x(t), \quad y(t) \qquad (t_1 \leq t \leq t_2)$$

having a constant length

$$J(C) = \int_{t_1}^{t_2} \sqrt{x'^2 + y'^2}\,dt = \text{constant}.$$

We shall proceed in a formal way to show that the solution is a circle. Let C be the solution, which we assume exists. Since J is constant we shall be interested only in variations δx, δy for which the first variation δJ of J along C is zero. Since I has a maximum value on C, it follows that the first variation δI of I along C vanishes whenever $\delta J = 0$. Because of linearity, this is possible only in case δI is a multiple of δJ, that is,

$$\delta I = \lambda \delta J \quad \text{or} \quad \delta(I - \lambda J) = 0.$$

Letting

$$F = \tfrac{1}{2}(x y' - y x') - \lambda (x'^2 + y'^2)^{1/2}$$

we see that the integral

$$I - \lambda J = \int_{t_1}^{t_2} F\, dt$$

has the property that its first variation $\delta(I - \lambda J)$ vanishes along C. It follows that C must satisfy the Euler equations

$$\frac{d}{dt} F_{x'} = F_x, \quad \frac{d}{dt} F_{y'} = F_y .$$

Consequently

$$(6{:}1) \quad -\frac{d}{dt}\left[\frac{y}{2} + \frac{\lambda x'}{(x'^2 + y'^2)^{1/2}}\right] = \frac{y'}{2}, \quad \frac{d}{dt}\left[\frac{x}{2} - \frac{\lambda y'}{(x'^2 + y'^2)^{1/2}}\right] = -\frac{x'}{2};$$

If arc length is chosen as the parameter, then

$$x'^2 + y'^2 = 1$$

and equations (6:1) take the simple form

$$-\lambda x'' = y', \quad \lambda y'' = x'.$$

Multiplying the first of these by y'/λ and the second by x'/λ we obtain by addition the relation

$$x' y'' - x'' y' = \frac{1}{\lambda} = \text{constant}.$$

The solution to our problem therefore must have constant curvature $1/\lambda$ and hence must be a circle of radius λ, as was to be proved.

The problem just described is a special case of a large class of problems, commonly called *isoperimetric problems*. These problems consist of maximizing or minimizing an integral I in a class of arcs along which certain integrals $J_1, \cdots, J_p$ take on fixed values. These problems may be in parametric or in non parametric form.

Perhaps the simplest isoperimetric problem is that of minimizing an integral

$$I(C) = \int_{x_1}^{x_2} f(x,y,y')\,dx$$

in the class of non parametric arcs

$$C: \qquad y(x) \qquad (x_1 \leqq x \leqq x_2)$$

which join two fixed points 1 and 2 in the xy-plane which satisfies a single isoperimetric condition

$$J(C) = \int_{x_1}^{x_2} g(x,y,y')\,dx = \text{constant}.$$

Except in singular cases a minimizing arc must satisfy the corresponding Euler equations

$$\frac{d}{dx} F_{y'} = F_y \tag{6:2}$$

where

$$F(x,y,y') = f(x,y,y') - \lambda g(x,y,y')$$

and λ is a suitably chosen constant. It should be observed that if one sought to minimize or maximize $J(C)$ holding $I(C)$ constant, then equations (6:2) would hold with $F = g - \mu f$. The resulting equation can be obtained by dividing (6:2), as it stands, by λ and setting $\mu = 1/\lambda$. This suggests the following.

Principle of Reciprocity. If C minimizes $I(C)$ subject to the condition $J(C)$ = constant, then normally C maximizes or minimizes $J(C)$ subject to the condition $I(C)$ = constant, according as the multiplier λ appearing in (6:2) is positive or negative.

It is not difficult to construct examples in which the principle of reciprocity fails to hold.

The equations (6:2) are the Euler equations of

$$I - \lambda J = \int_{x_1}^{x_2} F\,dx\,.$$

From this fact one should not conclude that C minimizes $I - \lambda J$. This can perhaps best be illustrated by the following example. It is a known fact that a hanging chain assumes a position so that its center of gravity is as low as possible. Thus, a perfectly flexible chain of constant density and of prescribed length can be used to define the solution (with the y-axis as the vertical axis) of the following minimum problem: To find among all arcs C in the xy-plane joining two fixed points 1 and 2 and having a given length

$$J(C) = \int_{x_1}^{x_2} (1 + y'^2)^{1/2} dx = \gamma$$

one which minimizes the integral

$$I(C) = \int_{x_1}^{x_2} y(1+y'^2)^{1/2}dx .$$

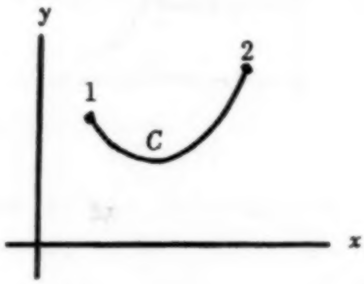

Fig. 6.1

As remarked above the solution C to this problem must satisfy the Euler equations of the integral

$$I - \lambda J = \int_{x_1}^{x_2} (y-\lambda)(1+y'^2)^{1/2}dx$$

where λ is a suitably chosen constant. The arc C is a catenary with the line $y = \lambda$ as its directrix. As was remarked in Section 2 for the case $\lambda = 0$, there may be two catenaries joining 1 and 2 having $y = \lambda$ as its directrix. The longer one fails to minimize $I - \lambda J$ but does minimize I in the class of arcs joining its end points and having the same length.

7. *Variable end point problems.* Hitherto we have been concerned mainly with the problem of minimizing an integral

$$I(C) = \int_{x_1}^{x_2} f(x,y,y')\,dx$$

in a class of arcs

C: $\qquad y(x) \qquad (x_1 \leqq x \leqq x_2)$

joining two fixed points 1 and 2. If now we permit the point 2 to be an arbitrary point on a prescribed curve

D: $\qquad x = X(u), \quad y = Y(u)$

as indicated in the figure, (next page), one encounters two new concepts, namely, transversality and focal points. In order to describe these concepts let us first recall certain facts concerning the problem of finding the shortest line joining a fixed point 1 to a point 2 on D. If C is such a line segment then it must have two properties:

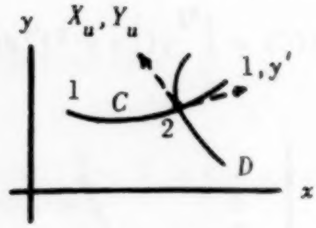

Fig. 7:1

(1) The line C is perpendicular to D at the point 2.

(2) The center of curvature of D at the point 2 cannot lie on C between 1 and 2.

One can convince himself of these facts by drawing diagrams.

In the calculus of variations the analogues of perpendicularity and centers of curvature are transversality and focal points. The condition of transversality can be stated as follows.

If an arc C minimizes $I(C)$ in the class of arcs joining the point 1 *to a point on a curve D, then at the point* 2 *of C on D, the condition*

$$(f - y' f_{y'}) X_u + f_{y'} Y_u = 0 \tag{7:1}$$

must hold, where 1, y' *and* X_u, Y_u *denote the vectors tangent to C and D respectively, as indicated in Figure* 7:1.

If $f = \sqrt{1 + y'^2}$, then this condition is equivalent to

$$X_u + y' Y_u = 0.$$

Consequently transversality becomes orthogonality in this case.

In order to establish (7:1) let

$$y(x, u) \qquad x_1 \leqq x \leqq X(u)$$

be a one parameter family of arcs joining the point 1 to the point $(X(u), Y(u))$ on D and containing C for $u = u_0$. Along this family the integral I has the value

$$I(u) = \int_{x_1}^{X(u)} f\big(x, y(x,u),\ y_x(x,u)\big)\, dx\, .$$

Since C is minimizing we have

$$I(u) \geqq I(u_0) = \text{value of } I \text{ along } C.$$

Conxequently $I'(u_0) = 0$, that is

$$I'(u_0) = fX_u \Big|^{x_2} + \int_{x_1}^{x_2} \{f_y\, y_u + f \quad y_{xu}\}\, dx = 0\,,$$

the arguments in f being those belonging ' , C. Since C minimizes $I(C)$ in the class of arcs joining its end points we have

$$\frac{d}{dx} f_{y'} = f_y$$

along C. Using this fact we obtain

$$I'(u_0) = fX_u \Big|^{x_2} + f_{y'}\, y_u \Big|_{x_1}^{x_2} = 0\,.$$

But

$$y(x_1, u) = y_1\,, \quad y(X(u),\ u) = Y(u)\,.$$

Differentiating with respect to u we see that

$$y_u = 0 \text{ at } x = x_1\,, \quad y_x X_u + y_u = Y_u \text{ at } x = x_2 = X(u_0)\,.$$

Consequently the above expression for $I'(u_0)$ becomes

$$I'(u_0) = [(f - y' f_{y'}) X_u + f_{y'} Y_u]^{x = x_2} = 0\,,$$

as was to be proved.

It remains to describe the generalization of the center of curvature. Recall that the locus of the centers of curvatures is the envelope of the normals of D. Thus, on a line C perpendicular to D the center of curvature is the point of contact of C with this envelope. This fact suggests that we replace the normals to D by the family of extremals transversal to D. This family will in general have an envelope E, as indicated in Figure 7:2. Consider an extremal C of this family.

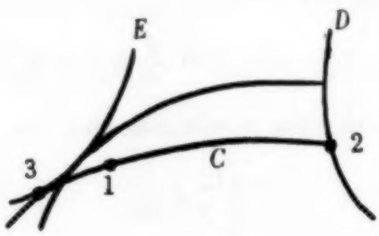

Fig. 7:2

The point 3 of contact of C with E is called a focal point of D. If the segment C_{12} of C is to be a solution to the problem here considered, the focal point 3 cannot lie between 1 and 2. The proof of this fact can be made by an extension of the argument made for the corresponding result for shortest distance problems.

8. *Minima of functions of integrals.* In the preceding pages we have been concerned with minima of single integral. The results obtained are readily extended to functions of integrals. To illustrate the method consider two integrals

$$I(C) = \int_{x_1}^{x_2} f(x,y,y')\,dx\,, \qquad J(C) = \int_{x_1}^{x_2} g(x,y,y')\,dx\,.$$

We seek to minimize the ratio

$$\frac{I(C)}{J(C)}$$

in the class of arcs

$$y(x) \qquad\qquad (x_1 \leqq x \leqq x_2)$$

joining two fixed points 1 and 2 in the xy-plane. Let C_0 be the minimizing arc. Then the first variation

$$\delta\frac{I}{J} = \frac{J\,\delta I - I\,\delta J}{J^2}$$

must vanish for all variations δy vanishing at x_1 and x_2. Setting

$$\lambda = \frac{I(C_0)}{J(C_0)} \tag{8:1}$$

we see that

$$\delta I - \lambda\delta J = \delta\int_{x_1}^{x_2} (f - \lambda g)\,dx = 0$$

along C_0 for all such variations δy. Consequently C_0 must satisfy the Euler equations

$$\frac{d}{dx}(f_{y'} - \lambda g_{y'}) = f_y - \lambda g_y\,.$$

This can be written in the form

$$\frac{d}{dx}f_{y'} - f_y = \lambda\left(\frac{d}{dx}g_{y'} - g_y\right).$$

In order to illustrate this result consider the problem of minimizing the ratio

$$\frac{\int_0^1 y'^2\,dx}{\int_0^1 y^2\,dx}$$

in the class of arcs $y(x)$ $(0 \leqq x \leqq 1)$ having

$$y(0) = 0, \quad y(1) = 0\,. \tag{8:3}$$

Here $f = y'^2$ and $g = y^2$. Consequently equation (8:2) takes the form

$$y'' + \lambda y = 0 .$$

The solutions of this equation satisfying (8:3) are of the form

$$y = \sin \sqrt{\lambda}\, x$$

where λ has one of the values π^2, $4\pi^2$, $9\pi^2$, $\cdots$, $n^2\pi^2$, $\cdots$. Moreover

$$\frac{\int_0^1 n^2\pi^2 \cos^2 n\pi x \, dx}{\int_0^1 \sin^2 n\pi x \, dx} = n^2\pi^2$$

as required by (8:1). This ratio has a minimum when $n = 1$.

It is clear that it is easy to set up problems which have no solution. For example the ratio

$$\frac{\int_{x_1}^{x_2} y \sqrt{1+y'^2}\, dx}{\int_{x_1}^{x_2} \sqrt{1+y'^2}\, dx}$$

represents the y-coordinate of the centroid of an arc C joining a point (x_1,y_1) to a point (x_2,y_2). Obviously this quantity has no minimum value in the class of all arcs joining these points. If one writes down the Euler equations corresponding to (8:2) the solution takes the form

$$y = \lambda + b \cosh \frac{x-a}{b} .$$

The line $y = \lambda$ is the directrix of this catenary and hence cannot be the y-coordinate of the centroid, as was required. The equations (8:1) and (8:2) for this problem accordingly have no solution.

University of California, Los Angeles, and
National Bureau of Standards.

CURRENT PAPERS AND BOOKS

Edited by

H. V. Craig

This department will present comments on papers previously published in the MATHEMATICS MAGAZINE, lists of new books, and book reviews.

In order that errors may be corrected, results extended, and interesting aspects further illuminated, comments on published papers in all departments are invited.

Communications intended for this department should be sent in duplicate to H. V. Craig, Department of Applied Mathematics, University of Texas, Austin 12, Texas.

First Year College Mathematics with Applications. By Paul H. Daus and William M. Whyburn. The Macmillan Company, New York, 1949, xiii + 495 pages.

This book is planned particularly for engineering and technical schools. Calculus and determinants are omitted. The chapter headings are as follows: Introduction; Linear Functions; Quadratic Function of One Variable; the Parabola; the Central Conics in Typical Forms; Further Study of Conics; Algebraic Functions; Exponential and Logarithmic Functions; Algebra of Trigonometric Functions; Analytic Geometry of Trigonometric Functions. There are tables of assorted kinds, a list of answers to problems, and a five-page index. The authors have included a ninety-day outline and a seventy-five-day outline, planned respectively for a class meeting three hours a week through two semesters and for a class meeting five hours a week through one semester. Oral and written problems appear. We find both elementary material (for review) "indicated by a dagger" and starred paragraphs and examples to be omitted if desired.

Proof-reading has been very carefully done. The reviewer sees exceptionally few errors. On page 25 in §4.12 we come upon the somewhat unusual word "equilibfrium" (lines 7,8); while at the end of the fifth line of this section "e+" should be replaced by "of". On page 367 in the first line of §52.1 for "protection" read "projection". On page 393 in §54.31 at the end of the page read "cos" for "ω".

It would appear better to write "ungraduated ruler" instead of "ruler" on page 365 (line 11). The name "rectangular hyperbola" might be given as well as "equilateral hyperbola". On page 267 in line 2 of the last paragraph of §41.1 "zero of $g(x)$" seems preferable to "root of $g(x)$". And of course, since the reviewer is an unreconstructed pure mathematician, teaching in a liberal arts college, she weeps for the Calculus in connection with graphs of polynomials, etc.

The authors are clear and accurate on one-to-one correspondence, on curve tracing of all sorts (in spite of the absence of the derivative), on systems of straight lines and of conics, in the introduction of

"$\sin^{-1} x$", in the discussion of parametric equations and the reasons for them, in the use of hyperbolic functions, in the introduction of occasional bits of history (i.e. the "three famous problems of antiquity"), and so on. Examples are good and plentiful, especially those involving Physics. The reviewer is cheered to find that most of analytic geometry as usually taught in a college of liberal arts is satisfactory for engineering and technical schools. Many types of problems are the same. At least, many of our favorite examples, including geometric theorems to be proved analytically, are found in this book. On page 355 appears the 12-ft. ladder resting "against a vertical wall and on the ground in such a way that it just clears a 3-ft. fence — placed 4 ft. from the wall", which has been previously brought to the reviewer by interested friends, including a college student and a milkman.

A few sentences are so good that they must be quoted. "It is more important to understand the complete analysis of any of these problems than to carry out all numerical calculations for some of them" (p. 151). "But the method is of more importance than the formula" (p. 170). "Any attempt at point-plotting requires the solutions of cubic equations with irrational roots, and this is discouraging, to say the least" (p. 278). "For purposes of prediction the use of an assumed law, without some evidence other than statistical data, is fraught with dangers and has many times led investigators to erroneous conclusions" (p. 314). Why not write that last sentence in letters of gold upon the walls where all would-be statisticians may read them!

Wellesley College. Marion E. Stark

Analytic Geometry. By W. A. Wilson and J. I. Tracey, Third Edition, D. C. Heath and Company, 1949.

The new edition like its predecessors covers the conventional topics of plane analytic geometry and provides a brief treatment of solid analytic geometry. The book contains a discussion of empirical equations and includes the method of moments as well as the method of averages in curve fitting. Polar coordinates are introduced in a separate chapter after considerable work has already been done with curves and equations in Cartesian coordinates. Spherical and cylindrical coordinates are introduced in the chapter on solid geometry.

The arrangement of the material seems very satisfactory. The reviewer would like to see a chapter on an introduction to the calculus preceding the chapters on higher plane curves and on tangents and normals. Such a chapter would add to the student's understanding of the higher plane curves and would make possible the elimination of some very specialized procedures in the study of tangents and normals.

The problem material is ample and varied. Problems of the numerical and drill type are numerous; problems of historical interest, and problems for the abler students are included. Much of the problem

material has not appeared in the earlier editions. Answers to most of the odd-numbered problems are given at the end of the book, while answers to the even-numbered problems are omitted. The student thus has both the convenience of a check on his work and the opportunity to learn to rely upon himself.

The book has the same clear, teachable presentation of the earlier editions, and a new and pleasing format.

Wellesley College. Helen G. Russell

College Algebra. By J. B. Rosenbach and E. A. Whitman, Third Edition, Ginn and Company, Boston, 1949. 508 pp. $3.00.

The *College Algebra* is designed to meet the needs of students whose preparation varies from two to four semesters of high-school algebra and of students whose preparation is several years in the past. To this end the early chapters contain a thorough review of elementary algebra. Both the earlier and the later chapters are arranged to permit considerable flexibility in the content of a college course. Such topics as complex numbers, theory of equations, permutations, combinations and probability, determinants, partial fractions, and infinite series are treated quite fully, but may be omitted without affecting the continuity of a course.

The book contains abundant problem material. There are many simple drill exercises following each principle discussed in the text, and there are also problems whose solution requires resourcefulness and real comprehension of fundamental principles. Answers are provided for odd-numbered problems, and are available in a separate pamphlet for all problems.

The presentation is clear, the illustrative examples well chosen. Occasionally the reviewer takes exception to the teaching methods employed, as, for instance, to the use of division rather than multiplication in the treatment of complex fractions, and to the special device for expanding a determinant of the third order. The "warnings" sometimes serve a useful purpose, but might often be omitted to allow the student to learn by experience rather than by advice. The historical and supplementary notes add material which enriches the teaching and may lead the student to further investigation of a particular topic.

The book has been written with the first-year college student in mind, and is of such size that he can easily carry it from class to class. Its more than 500 pages comprise a useful reference work for anyone whose knowledge of algebra needs refreshing.

Wellesley College. Helen G. Russell

PROBLEMS AND QUESTIONS

Edited by

C. W. Trigg, Los Angeles City College

Readers of this department are invited to submit for solution problems believed to be new and subject-matter questions that may arise in study, in research, or in extra-academic situations. Proposals should be accompanied by solutions, when available, and by such information as will assist the editor. Ordinarily, problems in well-known text-books should not be submitted.

In order to facilitate their consideration, solutions should be submitted on separate, signed sheets within three months after publication of the problems.

All manuscripts should be typewritten on 8½" by 11" paper, double-spaced and with margins at least one inch wide. Figures should be drawn in india ink and in exact size for reproduction.

Send all communications for this department to *C. W. Trigg, Los Angeles City College, 855 N. Vermont Ave., Los Angeles 29, Calif.*

PROPOSALS

63. *Proposed by Victor Thébault, Tennie, Sarthe, France.*

In the system of numeration having base 11, find a six-digit square of the form *abcabc*.

64. *Proposed by D. L. MacKay, Manchester Depot, Vt.*

Construct a trapezoid $ABCD$ given its diagonals and its non-parallel sides.

65. *Proposed by Howard Eves, Oregon State College.*

A thin disk of given radius is spinning on a fixed horizontal axis, and while spinning dips into a vat of liquid. Locate the position of the axis with respect to the surface of the liquid so that the spinning disk exhibits a maximum evaporation surface.

66. *Proposed by E. P. Starke, Rutgers University.*

Given four concylic points A, B, O, P such that $AO = BO$. Show that the ellipse through P having foci at A and B, and the circle through P having center at O are orthogonal or tangent according as O, P are on the same side of AB or on opposite sides.

67. *Proposed by N. A. Court, University of Oklahoma.*

Given two spheres (A), (B). Let p be the perpendicular dropped from B upon a plane (Q) passing through A. If the traces of p and (Q) upon the plane of the common circle (S), real or imaginary, of the two spheres are pole and polar with respect to (S), the two spheres are orthogonal.

68. *Proposed by P. A. Piza, San Juan, Puerto Rico.*

Show that it is possible to construct by ruler and compasses a Fermagoric triangle of order 3, i.e. a triangle ABC such that $AB^3 = AC^3 + BC^3$. In particular, let the line RH be divided into three segments such that $RG = s^2(r^2 + rs + 4s^2)$, $GF = (r^2 - s^2)(r^2 + rs + 4s^2)$, $FH = (r+s)(4s^3 \pm r^3)$, where r, s are positive integers with $s < r < s\sqrt[3]{4}$. Further, let RH be the diameter of a circle and let AB and CD be chords perpendicular to RH and passing through G and F respectively. Verify that this triangle ABC satisfies the above relation.

69. *Proposed by J. S. Miller, Dillard University.*

An automobile door stands open at right angles. The hinges are at the front of the door. If the automobile is given an acceleration a find the time in which the door closes.

Notice to Solvers.

"The really interesting thing about the solution of a problem is: How did the solver find his solution or construct his proof? The problems are seldom important in themselves, nor are the solutions any more important. The interest is in *finding* the solution — the heuristic aspect." This quotation is from a letter received from Kenneth May who offers below a heuristic treatment of problem 42. Readers are invited to follow his lead and to offer heuristic discussions in addition to formal solutions.

SOLUTIONS

Late Solutions

37, 38. Leo Moser, University of North Carolina.

The Four Basic Points of a Pencil of Conics.

2. [Sept. 1947] *Proposed by E. P. Starke, Rutgers University.*

What restrictions must be placed on the four basic points of a pencil of conics, if all the conics of the pencil are to have (a) a common center, (b) a common axis of symmetry?

Solution by the Proposer. We assume the four points to be distinct with no three collinear. Then the pencil of conics determined by them contains three degenerate conics consisting of pairs of lines through the four points. If the lines are intersecting, the degenerate conic has its center at the intersection. By the geometry of the figure no two degenerate conics of this type can have a common center. Therefore case (a) requires that two of the degenerate conics be pairs of parallel lines and hence that the four points be vertices of a parallelogram.

If each of the degenerate conics is a pair of intersecting lines we have already three non-collinear centers, which renders impossible a common axis of symmetry. Thus for case (b) at least one degenerate conic must be a pair of parallel lines so that the four points are vertices of a trapezoid. But an axis of symmetry of a pair of parallel lines is either the line midway between them or an arbitrary line perpendicular to them. Therefore, to satisfy (b), the line joining the intersection of the other two pairs of lines through the vertices of the trapezoid must be perpendicular to the parallel sides, and hence the four points form an isosceles trapezoid (with a rectangle as a special case.)

Conversely, if coordinate axes and constants a, b, c are so chosen that $(a+b,c)$, $(a-b,c)$, $(-a+b,c)$, $(-a-b,-c)$ are the vertices of an arbitrary parallelogram then the pencil of conics is given by

$$m[(by-cx)^2-a^2c^2]+n(y^2-c^2)=0,$$

with m and n arbitrary. Clearly there is a common center, (0,0), satisfying (a).

If the points are the vertices of the isosceles trapezoid $(\pm a,0)$, $(\pm b,c)$, then the pencil becomes

$$m[(ay-by-ac)^2-c^2x^2]+ny(y-c)=0.$$

Here every conic is symmetric to the y-axis, satisfying (b).

Also solved by Howard Eves, Oregon State College.

A Right Smart Census Taker.

42. [Nov. 1949] *Proposed by V. C. Throckmorton, Los Angeles City College.*

Encountering a man on the porch of his home, a census taker asked, "What are the ages of the persons living here?" The man replied, "My age is the sum of the ages of my wife, son and daughter. Each of our ages is a square number. My father's age is the sum of my age and the ages of my wife and daughter. Although he has passed the prime of life, his age is a prime number." What ages did the unstartled census taker record, and what obvious remark did he make about the wife's age?

I. *Solution by Monte Dernham, San Francisco.* Here is how that census taker figured. If the daughter, son, wife, man and pater are respectively d^2, s^2, w^2, m^2 and p years of age, then

$$d^2+s^2+w^2=m^2;$$

again,

$$d^2+w^2+m^2=p;$$

whence

$$2(d^2+w^2)+s^2=p.$$

Since p, a prime >2, is odd, s^2 is odd. Now, if $s^2\geqq 25$, then $w^2\geqq 36$, $m^2>64$ and $p>101$, a highly unrealistic situation, which becomes increasingly impossible as the *ante* is raised. Therefore $s^2=1$ or 9, and

$$w^2+d^2=m^2-1 \quad \text{or} \quad m^2-9.$$

Thereupon the census taker jotted down the following possibilities, none other being within admissible bounds:

m^2	64	49	36	25	16
m^2-1	63	48	35	24	15
m^2-9	55	40	27	16	

By some curious stroke of good luck, it happened that the census taker had just been reading Chapter 8 of Oystein Ore's *Number Theory and its History,* and thus could see at a glance that one, and only one,

of the nine numbers in the second and third rows could be written as the sum of two squares, *viz.*, 40. Unhesitatingly he recorded the five ages, remarking at the same time that the daughter's age followed by the son's age would denote the man's age, but if a dot were to be placed between the two digits, like this: 4·9, why then 'twould be the wife's age — obviously, a perfect 36! So the pater was 89.

II. *Heuristic Discussion by Kenneth May, Carleton College, Northfield, Minnesota.* The problem is trivial in the sense that the solution emerges from a few trials with the list of small perfect squares. However, the writer was interested in seeing how people would actually go about finding the solution and so submitted the problem to his freshman mathematics class. Students found it more difficult to explain their procedures than to find the answer, but the following three methods could be singled out: (1) Starting with 25, consider successive squares as possible ages of the father and see whether any one is the sum of three preceding ones. The first one with this property yields a prime grandfather's age and hence a solution. (*Clyde Roberson*). (2) Start with pairs of small squares and see if any have a sum equal to the difference of two following perfect squares. (*Marlene Erickson, Harold Klepfer*). (3) The conditions of the problem, including the fact that the census-taker was "unstartled" make it "fairly safe" to guess that the parents are at least 25 and not more than 49. Since they must be unequal, we consider the possibilities 49,36; 36,25; 49,25. The first one yields the solution. (*Stephan West*). No student suggested the following line of reasoning: (4) Since the census taker made an "obvious remark" about the mother's age, it must have some simple numerical property not already mentioned. Since it is already known to be a perfect square related by sums to the other ages, could this additional property involve some simple relation in terms of products? Possibly her age is the product of her childrens' ages. This suggests 36 which is the smallest perfect square equal to the product of two perfect squares.

It is interesting to observe how these methods make increasing use of the given information in order to reduce the number of trials. Method (4) amounts to guessing the solution and then testing it. Such techniques are intriguing, but they have the disadvantage of being highly special to the particular problem and give us no clue to such important questions as (A) Have we found all the solutions? (B) Are the conditions of the problem insufficient or redundant? (C) What are the mathematical principles which underlie the problem? The method (2) is very suggestive of answers to these questions. Since we are looking for two perfect squares whose sum is the difference of two perfect squares, we may recall that the differences of successive perfect squares are the successive odd numbers. It follows that any two squares whose sum is odd will have the desired property. In fact if $x^2 + y^2$ is odd, $x^2 + y^2 + z^2 = w$, where $z = (x^2 + y^2 - 1)/2$ and $w = z + 1$.

It follows immediately that there are an infinity of perfect squares satisfying this condition of the problem. Those which give possible ages for people are 1,4,4,9; 1,16,64,81; 4,9,36,49. The first and third yield prime grandfather's ages, but the first is certainly "startling", and so is the second. Noting that the differences of alternate perfect squares are successive multiples of 4, we may look for two squares whose sum is a multiple of four. This yields 4,16,16,36, both startling and giving a non-prime grandfather's age. It is easy to see that no further unstartling ages can be found by considering differences of more widely separated squares. We can now answer the above questions as follows: (A) Yes, there is only one solution, but there are an infinity of perfect squares satisfying the equation $x^2 + y^2 + z^2 = w^2$. (B) The condition of primality is redundant. The solution of the problem is determined by the sum relations and the fact that the census-taker was "unstartled". It would be better to omit this condition and to indicate that the census taker made an "obvious" remark also about the grandfather's age! (C) The problem is connected with the relations between perfect squares and arithmetic progressions.

Also solved by H. E. Bowie, American International College; B. K. Gold, Los Angeles City College; P. Na Nagara, College of Agriculture, Bangkhen, Bangkok, Thailand; D. B. Pheley, Los Angeles City College; W. R. Ransom, Tufts College; L. A. Ringenberg, Eastern Illinois State College; Fern Spivey, Palms Junior H. S., Los Angeles; and W. R. Talbot, Jefferson City, Mo.

Condition for Concurrency of h_a, m_b, t_c

43. [Nov. 1949] *Proposed by Victor Thébault, Tennie, Sarthe, France.*

Show that the necessary and sufficient condition that the altitude AA', the median BB', and one of the bisectors of the angle C of a triangle ABC be concurrent is that $\sin A/\cos B = \pm \tan C$.

I. *Solution by P. Na Nagara, College of Agriculture, Bangkhen, Bangkok, Thailand.* $\sin A/\cos B = BC \sin B/AC \cos B = (BC/AC) \tan B = (BC/AC)/(AA'/BA')$, and $\tan C = AA'/A'C$, so the given condition is equivalent to $BA'/A'C = BC/AC = BC/2(B'C)$.

Let AA' cut BB' at O, and let D lie on BB' extended so that $OB' = B'D$. Since $AB' = B'C$, $AOCD$ is a parallelogram, so AA' is parallel to DC. Then $BA'/A'C = BO/OD = BO/2(OB')$. If follows that if the condition holds, $BC/B'C = BO/OB'$, so CO bisects angle ACB and the three lines are concurrent. Conversely, if CC' passes through O, then $BC/B'C = BO/OB'$, so $BC/AC = BO/OD = BA'/A'C$.

If CC' is an exterior bisector, a like procedure may be employed, with $\tan C = -AA'/A'C$ and D falling on BB'.

II. *Solution by D. L. MacKay, Manchester Depot, Vt.* By Ceva's Theorem and its converse, the necessary and sufficient condition that the three lines be concurrent is $(A'B)(B'C)(C'A) = (A'C)(B'A)(C'B)$ or $(c \cos B)(b/2)(bc/[a+b]) = (\pm b \cos C)(b/2)(ac/[a+b])$. Whence

$\cos B = \pm(a\cos C)/c = \pm\sin A\cos C/\sin C$. Therefore $\sin A/\cos B = \pm\tan C$, depending upon whether angle C is acute or obtuse.

III. *Solution by B. K. Gold, Los Angeles City College.* If the vertices of the triangle are $A(b,c)$, $B(0,0)$, and $C(a,0)$, then we have the equations

$$\begin{aligned} &\text{altitude } AA': && x-b=0,\\ &\text{median } BB': && cx-(a+b)y=0, \text{ and}\\ &\text{bisector } CC': && cx+[a-b\pm\sqrt{c^2+(a-b)^2}]y-ac=0. \end{aligned}$$

If these lines are concurrent then the determinant of the coefficients vanishes, so $a^2-ab\pm b\sqrt{c^2+(a-b)^2}=0$ or

$$[ac/\sqrt{b^2+c^2}\sqrt{c^2+(a-b)^2}]/[b/\sqrt{b^2+c^2}] = \pm c/(a-b).$$

Hence $\sin A/\cos B = \pm\tan C$.

IV. *Solution by L. M. Kelly, Michigan State College.* It is well known that the barycentric coordinates of A', B', C', are respectively $(0,\tan B,\tan C)$, $(1,0,1)$, $(\sin A,\sin B,0)$. Thus the equations of the three lines are $x_2\tan C-x_3\tan B=0$, $x_1-x_3=0$, and $x_1\sin B-x_2\sin A=0$. The necessary and sufficient condition that these meet in a point is

$$\begin{vmatrix} 0 & \tan C & -\tan B\\ 1 & 0 & -1\\ \sin B & -\sin A & 0 \end{vmatrix} = 0.$$

It follows that $\tan C = \sin A/\cos B$. For an external bisector, CC', the left hand member would turn out to have a negative sign.

Also solved by R. E. Horton, Los Angeles City College, and W. R. Talbot, Jefferson City, Missouri.

Editorial Note: The construction of ABC having AA', BB' and the external angle bisector concurrent is treated in *National Mathematics Magazine*, **15**, 149, (Dec. 1940); **18**, 91, (Nov. 1943). The case of the interior bisector is dealt with in *The American Mathematical Monthly*, **44**, 599, (1937); **47**, 176, (1940) where the necessary and sufficient condition appears in the form $c^2 = b^2+a^2-2ba^2/(b\pm a)$.

Pythagorean Triangles with Area Equal to Perimeter

44. [Nov. 1949] *Proposed by M. T. Goodrich, Keene Teachers College, Keene, N. H.*

Find all right triangles such that the sides are integers and the perimeter is numerically equal to the area.

I. *Solution by H. E. Bowie, American International College.* Let c be the hypotenuse and a and b the other two sides. Then

$$a^2+b^2=c^2 \text{ and } a+b+c=ab/2$$

Eliminating c from these two equations, we have $a = 4 + 8/(b-4)$. Now $8/(b-4)$ is a positive integer for $b = 5$, 6, 8 and 12 only. Since the equations are symmetrical in a and b, the only two distinct solutions are $a = 5$, $b = 12$, $c = 13$ and $a = 6$, $b = 8$, $c = 10$.

II. *Solution by Louis de Branges, III, Student, Massachusetts Institute of Technology.* The parametric equations for a Pythagorean triangle are

$$a = (x^2 - y^2)z, \quad b = 2xyz, \quad c = (x^2 + y^2)z,$$

where x, y, z are positive integers with $x > y$. The conditions of the problem are met if $2xz(x+y) = xyz^2(x+y)(x-y)$. Dividing by the non-zero factors, we have $x = y + 2/yz$. Thus (y, z) is restricted to $(1,1)$, $(1,2)$ and $(2,1)$. The first two pairs result in the same triangle with sides 6, 8, 10. The second pair gives the only other solution, a triangle with sides 5, 12, 13.

Also solved by Louis Berkofsky, Watertown, Mass.; M. P. de Regt, Walnut Creek, Calif.; Monte Dernham, San Francisco; Howard Eves, Oregon State College; B. K. Gold, Los Angeles City College; Thomas Griselle, Hollywood, Calif.; D. L. MacKay, Manchester Depot, Vt.; P. Na Nagara, Bangkok, Thailand (2 ways); W. R. Ransom, Tufts College; L. A. Ringenberg, Eastern Illinois State College; E. D. Schell, Washington, D. C.; W. R. Talbot, Jefferson City, Mo.; P. D. Thomas, Washington, D. C.; W. I. Thompson, Los Angeles City College; Samuel Wolf, Wilkinsburg, Pa.; and the proposer (2 ways).

Editorial Note: If the problem merely required that the legs be integers, the result would have been the same, as may be seen from Bowie's solution. Berkofsky showed that the 3, 4, 5 triangle is the only other triangle with a perimeter that is an integral multiple (m) of the area, and in that case $m = 2$.

This problem has enjoyed some popularity in the past, see, e.g.: *Ladies' Diary*, 34, (1828); *The Lady's and Gentleman's Diary*, London, 49-50, (1865); Dickson, *History of the Theory of Numbers*, Stechert (1934), Vol. II, 180, 195, 199; Alliston, *Mathematical Snack Bar*, Heffer (1936), 1; Kraitchik, *Mathematical Recreations*, Norton (1942), 101. Whitworth and Biddle [*Math. Quest. Educational Times*, 5, 54-6, 62-3, (1904)] proved that there are only five Heronian triangles whose area equals the perimeter, namely: (5,12,13), (6,8,10), (6,25,29), (7,15,20), (9,10,17).

QUICKIES

From time to time as space permits this department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solution and the source, if known.

Q 6. There are n players in an elimination-type singles tennis tournament. How many matches must be played (or defaulted) to determine the winner? [Submitted by Fred Marer].

Q 7. Find the first derivative of $y = nx^{n/(a \ln x)}$. [Submitted by Samuel Skolnik].

Q 8. Five cards are drawn at random from a pack of cards which have been numbered consecutively from 1 to 97, and thoroughly shuffled. What is the probability that the numbers on the cards as drawn are in increasing order of magnitude? [Submitted by J. R. Ziegler].

Q 9. From each of two diagonally opposite corners of an 8″ x 8″ board a 1″ x 1″ square is cut. Can the remainder of the board be completely covered by 2″ x 1″ strips without overlapping? [Submitted by Raphael Robinson].

Q 10. Around a cylindrical tube, outside circumference 4 inches, length 9 inches, ten turns of a wire are helically wrapped. The ends of the wire coincide with the ends of the same cylindrical element. Find the length of the wire.

Q 11. Find the prime factors of 1,000,027.

ANSWERS

A 6. Each match has one loser, each loser loses only once, so there are $n-1$ losers, hence $n-1$ matches.

A 7. Since $x = e^{\ln x}$, $y = ne^{n/a}$, so $dy/dx = 0$.

A 8. The permutations of any five numbers are 5! so the probability is 1/120.

A 9. No. Consider the original board divided into 64 equal squares, which have been alternately colored red and black. The cuts remove two squares having the same color. No matter how oriented, each 2″ x 1″ strip will cover one red and one black square, so any arrangement of the strips will leave two squares of the same color uncovered.

A 10. Roll the cylindrical surface (and wire) onto a plane. The element (9 inches), the repeated circumference (10 x 4 inches), and the wire (L) now form a right triangle. Hence, $L = \sqrt{81 + 1600} = 41$ inches.

A 11. $1{,}000{,}027 = (100)^3 + (3)^3 = (100 + 3)(10{,}000 - 300 + 9) = (103)(9709) = (103)(7)(1387) = (103)(7)(1460 - 73) = (103)(7)(73)(19)$.

THE CHRONOLOGY OF PI

Herman C. Schepler

(Continued from the March-April issue)

1872. Augustus De Morgan (1806-1871), born in Madras, was a severe critic of the would-be circle squarer. Educated at Trinity College, Cambridge, and became professor at the University of London in 1828; celebrated teacher who also contributed to algebra and the theory of probability. Most of his references to circle squaring appear in his *Budget of Paradoxes* which was edited and published by his wife in 1872, after his death. Among De Morgan's many antics regarding the circle-squarer, he ordained St. Vitus as the patron saint of the circle-squarer and suggested that Mr. James Smith (see 1860, James Smith) was seized with the *morbus cyclometricus,* defined as the *circle-squaring disease.* He tells of many of his experiences with cyclometers, such as the Jesuit who came from South America in about 1844, bringing a quadrature and a newspaper clipping announcing that a reward was ready for the discovery in England. [5], 330; [7]; [8], 305.

1873. William Shanks, (1812-1882) (England). 707 places, using Machin's formula. English mathematician. (See 1853, William Shanks). The value of π to 707 places is:

3.14159 26535 89793 23846 26433 83279 50288 41971 69399 37510 58209 74944
59230 78164 06286 20899 86280 34825 34211 70679 82148 08651 32823 06647
09384 46095 50582 23172 53594 08128 48111 74502 84102 70193 85211 05559
64462 29489 54930 38196 44288 10975 66593 34461 28475 64823 37867 83165
27120 19091 45648 56692 34603 48610 45432 66482 13393 60726 02491 41273
72458 70066 06315 58817 48815 20920 96282 92540 91715 36436 78925 90360
01133 05305 48820 46652 13841 46951 94151 16094 33057 27036 57595 91953
09218 61173 81932 61179 31051 18548 07446 23798 34749 56735 18857 52724
89122 79381 83011 94912 98336 73362 44193 66430 86021 39501 60924 48077
23094 36285 53096 62027 55693 97986 95022 24749 96206 07497 03041 23668
86199 51100 89202 38377 02131 41694 11902 98858 25446 81639 79990 46597
00081 70029 63123 77381 34208 41307 91451 18398 05709 85 &c.

[5], 206; [11], 261; [12], 77; [18], 16; [20], 188.

1876. Alick Carrick (England). 3 1/7 (3.1428571 ···). Proposed in his book, *The Secret of the Circle, Its Area Ascertained.* Nature 15: 1876.

1879. Pliny E. Chase, LLD (Haverford, Penna.) 3.14158499 ··· by geometrical construction. Pamphlet, *Approximate Quadrature of the*

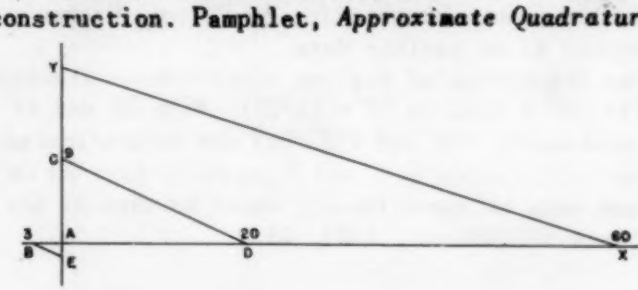

Fig. 4

Circle, Haverford, Penn., June 16, 1879. On the rectangular coordinates X, Y, lay off from a scale of equal parts, $AB = 3$; $AC = 9$; $AD = 20$; and $AX = 60$. Join C and D and through B draw BE parallel to CD intersecting the Y-axis at E. Take $EY=AD$, and join X and Y. Then $XY:AD$::Circum:Diam., nearly. The error is $-.00000766 \cdots$. [9], 9.

1882. Ferdinand Lindemann (1852-1939) (Germany) proved π to be transcendental, i.e., cannot be represented as the root of any algebraic equation with rational coefficients. He also proved the quadrature impossible. [5], 446; [19], 41.

1888. Sylvester Clark Gould (1840-1909). U.S.A. Editor of *Notes and Queries*, Manchester, New Hampshire. Due to popular demand for information on the subject of the quadrature, Gould was provoked to compile a bibliography on the subject which he titled, *What is The Value of Pi*, and published in 1888. 100 titles are presented which give the results of 63 writers. With the exception of two items the titles are all dated in the 19th century. [9].

1892. New York Tribune. (U.S.A.). 3.2 A writer announced this ratio as the rediscovery of a long lost secret which consisted in the knowledge of a certain "Nicomedean line". This announcement caused considerable discussion, and even near the beginning of the Twentieth Century 3.2 had its advocates as against the accepted ratio $3.14159 \cdots$. [18], 29.

1906. A. C. Orr. (U.S.A.). Literary Digest, Vol. 32. A mnemonic sentence is given for π. The number of letters in the words give the digits in the number π to 30 places.

> "Now I, even I, would celebrate
> In rhymes inapt, the great
> Immortal Syracusan, rivaled nevermore,
> Who in his wondrous lore,
> Passed on before
> Left men his guidance
> How to circles mensurate".

A sentence giving π to 31 places appears in [14], 67; a French verse to 30 places may be found in [16], 89. [6], 19; [15], 50.

1913. E. W. Hobson (England). $3.14164079 \cdots$. By a geometrical construction. Given in reference [10], by Hobson. Probably developed by another author at an earlier date.

Let r be the radius of a given circle whose diameter is AOB. Let $OD = (3/5)r$; $OF = (3/2)r$; $OE = (1/2)r$. With DE and AF as diameters, describe semicircles DGE and AHF. Let the perpendicular to AB through O cut these semicircles in G and H respectively. GH is the side of a square whose area is approximately equal to that of the given circle. The error is $+.0000481 \cdots$. [10], 35.

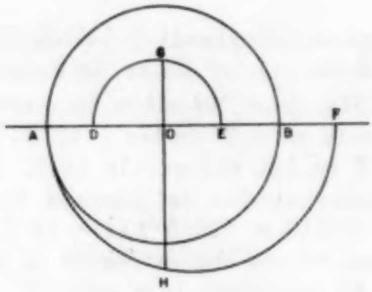

Fig. 5

1913. Srinivasa Ramanujan (Ramanujacharya) (1887-1920) (India) 3.1415926525826 ⋯. A mathematician who contributed several notweorthy approximations for evaluating π, both in empirical formulae and geometrical constructions. See *Approximate Geometrical Constructions for Pi*, Quarterly Journal of Math., *XLV*, 1914, pp. 350-374. See also, *Collected Papers of Srinivasa Ramanujacharya*, edited by G. H. Hardy, P. V. Seshu Aiyer and B. M. Wilson, published by University Press, Cambridge, England, 1927. A curious approximation to π obtained empiracally is:

$$\left[9^2 + \frac{(19)^2}{22}\right]^{\frac{1}{4}} = 3.1415926525826 \cdots.$$

The following construction is given for this value: Let *AB*, Figure 6, be a diameter of a circle whose center is *O*. Bisect the arc *ACB* at *C* and trisect *AO* at *T*. Draw *BC* and on it lay off *CM* and *MN* equal to *AT*. Draw *AM* and *AN* and on *AN* from *A*, lay off *AP* equal to *AM*. Through *P* draw *PQ* parallel to *MN* and meeting *AM* at *Q*. Draw *OQ* and through *T* draw *TR* parallel to *OQ*, and meeting *AQ* at *R*. Draw *AS* perpendicular to *AO* and equal to *AR*, and draw *OS*. Then the mean proportional between *OS* and *OB* will be very nearly equal to a sixth of the circumference, the error being less than 1/12 inch when the diameter is 8000 miles. The error is −.0000000010072 ⋯.

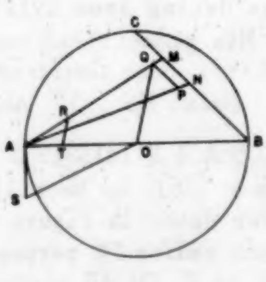

Fig. 6

1914. T. M. P. Hughes. (England) 3.14159292035 ··· by geometrical construction. *The Diameter of a Circle Equal in Area to any Given Square,* Nature, *93*:1914, page 110. This is a construction for the ratio 355/113. Make a circle with diameter = 11.3. Let $AX = 8\ 7/8$. Draw a perpendicular from X to cut the circle in Y. Join AY, and extend the line to Z. The construction is derived from the approximation: $AX:AB = \pi/4 = 355/4\cdot 113 = 710/8\cdot 113 = (8\ 7/8)/11.3$. Any circle with its diameter upon AB and one extremity of that diameter at A, will cut the line AZ (or AZ produced) in a point Y', making AY' the side of a square equal in area to the circle. Also, a line from Y' perpendicular to AB will cut the diameter in a point X', making AX' equal to 1/4 the circumference of the circle. Again, any square with its base upon AB and a corner at A, will cut AZ with side $X'W$ in a point Y', making AY' the diameter of an equal circle. The error of this quadrature is −.000000266 ···.

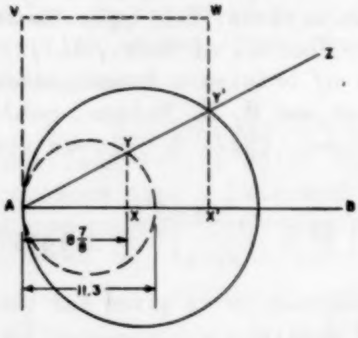

Fig. 7

1914. Scientific American, Mar. 21, 1914. (U.S.A.). The following mnemonic sentence is given for pi: "See I have a rhyme assisting My feeble brain its tasks sometimes resisting". The number of letters in the words give the digits in the number π to twelve places.

1928. Gottfried Lenzer. U.S.A. 3.1378 ···, an approximation obtained geometrically. In March 1928, Lenzer bequethed the University of Minnesota with 60 drawings dating from 1911 to 1927, regarding the three classical problems. His geometrical construction for squaring the circle gave $\pi = 3.1378\cdots$. *Some Constructions for the Classical Problems of Geometry* Amer. Math. Mo., *37*, Aug.-Sept., 1930. p. 343.

1933. Helen A. Merrill (U.S.A.) 3.141591953 ··· by geometrical construction. Given in reference [16], by Merrill. Probably developed by another author at an earlier date. In Figure 8, let AB, the diameter of a circle, be 1. Construct radius OE perpendicular to it. Tangents drawn at A and E intersect at F. On AB produced, lay off $BC = 1/10$, and $BD = 2/10$. At D draw DH perpendicular to AD and meeting FE extended

at H. Join A and H. From F on FA extended, lay off FG equal to FC. Through G draw a line parallel to AH, meeting FH extended in P. Solving from the relationships set up in the construction, $GP = 3.141591953 \cdots$. The error is $-.000000700 \cdots$. [16], 86.

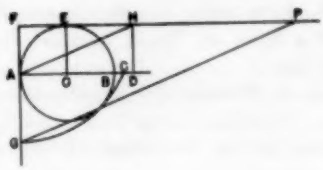

Fig. 8

1934. Carl Theadore Heisel (U.S.A.). 3 13/81 (3.1604938 ⋯). Proposed in his book, *Mathematical and Geometrical Demonstrations*. This book is 1¼" thick by 6½" x 10". It was printed at the author's expense and distributed to libraries without charge.

1941. Miff Butler. U.S.A. *Geomath*, General Engineering Co., Casper, Wyoming, 1941. Claimed discovery of a new relationship between π and e. Stated his work to be the first basic mathematical principle ever developed in the U.S.A. He got his congressman to read it into the Congressional Record on June 5, 1940.

1949. U.S. Army (U.S.A.). 2,035 places. Yielding to an irresistible temptation, some mathematical machine operators presented the problem of evaluating π to Eniac, the all-electronic calculator at the Army's Ballistic Research Laboratories in Aberdeen, Maryland. The machine's 18,800 electron tubes went into action and computed π to 2,035 places in about 70 hours. In 1873, William Shanks gave the value of π to 707 decimal places (527 correct). The computation took him more than 15 years. Scientific American, Dec., 1949, p. 30 and Feb., 1950, p. 2.

Stanford Research Institute,
Stanford, California.

MATHEMATICS MAGAZINE

Index to Volume XXIII September 1949 - June 1950

Articles - by Titles

Page

Articles - by Authors

The digit after the author's name indicates the number of this volume in which his biographical sketch appears. When the digit is missing, the sketch was published in a previous volume.

Current Papers and Books

Problems and Questions

Contributors

Numbers refer to pages. Bold type indicates a problem solved and solution published; italics, a problem solved but solution not published; ordinary type, a problem proposed.

Solutions

The page on which a solution appears is in parentheses following the number of the problem.

OUR CONTRIBUTORS

Richard F. Arens, Assistant Professor of Mathematics, University of California, Los Angeles, was born in Iserlohn, Germany in 1919. A graduate of University of California, Los Angeles (A.B.'41) and Harvard (A.M.'42, Ph.D.'45), Dr. Arens was a teaching fellow and tutor at Harvard, 1942-45 and then served as a staff assistant at the Institute for Advanced Study before being appointed to his present position. His special mathematical interests include topology and topological algebra, functional analysis, and harmonic analysis on groups.

Rodney T. Hood, Instructor, Beloit College, attended Oberlin College (A.B.'46) and the University of Wisconsin (M.A.'47), and is at present working on his Ph.D. degree at the latter institution. He is also interested in religion, and plans to prepare for the ministry.

Murray R. Spiegel, of the Rensselaer Polytechnic Institute, graduated from Brooklyn College (A.B.'43) and Cornell University (M.S.'46, Ph.D.'49). He has taught at Harvard, Cornell and Columbia and also worked as an applied mathematician for the Wave Propagation Group at Columbia during the war and later as a physicist at Oak Ridge.

Robert E. Horton, Instructor, Los Angeles City College, attended the University of California, Los Angeles (A.B.'36, M.A.'47) and the University of Southern California. During the war he served as Director of Ground Training at an Air Force school and later with the 15th Air Force in Italy. He has also taught at the University of California, Los Angeles. His chief mathematical interests are in geometry and analysis.

Magnus R. Hestenes, Professor of Mathematics, University of California, Los Angeles, was born at Bricelyn, Minnesota in 1906, and attended St. Olaf College (A.B.'27), University of Wisconsin (A.M.'28) and the University of Chicago (Ph.D.'32). A National Research Fellow and Pierce Instructor at Harvard (1933-36) he next taught for a year at the University of California, Los Angeles, and then returned to Chicago as Assistant and later Associate Professor. He was appointed to his present position in 1947. During the war he worked with the Columbia applied mathematics group (Associate Director 1944-45). Well known for his researches in the calculus of variations, Professor Hestenes has served on the Council of the American Mathematical Society and as a representative of the Society to the National Research Council. He has contributed especially to the problem of Bolza and the general problem of Mayer in the calculus of variations.